

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:22:20 ; Search time 75 Seconds
(without alignments)
1188.596 Million cell updates/sec

Title: US-10-022-609-11

Perfect score: 3470

Sequence: 1 ARAPQRGSLSPSRDKLFPN.....RLADSRNPAGRFSTQEEIQ 669

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3470	100.0	669	22	Human Heregulin HR
2	3470	100.0	669	21	Human Heregulin, H
3	3466	99.9	669	13	Human heregulin-al
4	3283.5	94.6	675	13	Human heregulin-be
5	3283.5	94.6	675	19	Amino acid sequenc
6	3283.5	94.6	675	21	Human Heregulin HR
7	3283.5	94.6	675	21	Human Heregulin HR
8	3283.5	94.6	675	21	Human Heregulin HR
9	3283.5	94.6	675	21	Human Heregulin HR
10	3283.5	94.6	675	22	Human Heregulin, H

11	3282.5	94.6	675	21	AAV71202	Human Heregulin HR
12	3281.5	94.6	675	19	AAW74493	Amino acid sequenc
13	3280.5	94.5	675	21	AAV71188	Human Heregulin HR
14	3280.5	94.5	675	21	AAV71189	Human Heregulin HR
15	3280.5	94.5	675	21	AAV71191	Human Heregulin HR
16	3280.5	94.5	675	21	AAV71198	Human Heregulin HR
17	3279.5	94.5	675	21	AAV71182	Human Heregulin HR
18	3279.5	94.5	675	21	AAV71194	Human Heregulin HR
19	3278.5	94.5	675	21	AAV71180	Human Heregulin HR
20	3278.5	94.5	675	21	AAV71185	Human Heregulin HR
21	3278.5	94.5	675	21	AAV71197	Human Heregulin HR
22	3277.5	94.5	675	21	AAV71181	Human Heregulin HR
23	3277.5	94.5	675	21	AAV71183	Human Heregulin HR
24	3277.5	94.5	675	21	AAV71184	Human Heregulin HR
25	3277.5	94.5	675	21	AAV71186	Human Heregulin HR
26	3277.5	94.5	675	21	AAV71187	Human Heregulin HR
27	3277.5	94.5	675	21	AAV71190	Human Heregulin HR
28	3277.5	94.5	675	21	AAV71201	Human Heregulin HR
29	3276.5	94.4	675	21	AAV71178	Human Heregulin HR
30	3276.5	94.4	675	21	AAV71192	Human Heregulin HR
31	3276.5	94.4	675	21	AAV71195	Human Heregulin HR
32	3275.5	94.4	675	21	AAV71196	Human Heregulin HR
33	3274.5	94.4	675	21	AAV71179	Human Heregulin HR
34	3274.5	94.4	675	21	AAV71193	Human Heregulin HR
35	3268.5	94.2	675	19	AAW74509	Amino acid sequenc
36	3267.5	94.0	675	19	AAW74503	Amino acid sequenc
37	3262.5	94.0	675	19	AAW74488	Amino acid sequenc
38	3259.5	93.9	675	19	AAW74500	Amino acid sequenc
39	3252.5	93.7	675	19	AAW74487	Amino acid sequenc
40	3252.5	93.7	675	19	AAW74496	Amino acid sequenc
41	3252.5	93.7	675	19	AAW74491	Amino acid sequenc
42	3252.5	93.7	675	19	AAW74489	Amino acid sequenc
43	3249.5	93.7	675	19	AAW74490	Amino acid sequenc
44	3249.5	93.6	675	19	AAW74486	Amino acid sequenc
45	3247.5	93.6	675	19	AAW74492	Amino acid sequenc

ALIGNMENTS

RESULT 1				
AAV71171	AAV71171 standard; Protein; 569 AA.			
ID	AAV71171	standard; Protein; 569 AA.		
XX	AC	AAV71171		
XX	DT	21-SEP-2000 (first entry)		
XX	DE	Human Heregulin HRG-alpha protein from ORF3.		
XX	KW	Heregulin; HRG alpha; human; hair cell; inner-ear-supporting cell;		
XX	KW	activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;		
XX	KW	hair cell related hearing disorder; ototoxic injury; tissue damage;		
XX	KW	acoustic assault; degenerative hearing loss; balance impairment;		
XX	OS	treatment; surgical injury; physical injury; inner ear disorder.		
XX	XX	Homo sapiens.		
XX	Key	Location/Qualifiers		
FT	Cleavage-site	2..3	"Protease processing site"	
FT	Cleavage-site	/note= 8..9	"Protease processing site"	
FT	Cleavage-site	23..24	"Protease processing site"	
FT	Cleavage-site	/note= 33..34	"Protease processing site"	
FT	Cleavage-site	/note= 36..37	"Protease processing site"	
FT	Modified-site	/note= 42..43	"Protease processing site"	
FT	Misc-difference	45	"Glycosaminoglycan addition site"	
FT		/label= Initiator_methionine		

Db 661 RFTQEEIQ 669

RESULT 2
AAU09886
ID AAU09886 standard; Protein; 669 AA.
AC AAU09886;
XX
XX 26-FEB-2002 (first entry)
XX Human heregulin, HRG-alpha.
XX
KW Human; heregulin; antiasthmatic; antiinflammatory; vulnery; antitumor;
epithelial cell growth; HER2; HER3; HER4; heregulin; lung cell; asthma;
lung surfactant protein A; respiratory distress; emphysema; pneumonitis;
chronic obstructive pulmonary disease; COPD; chronic bronchitis; ulcer;
neonatal pulmonary disease; neonatal respiratory distress syndrome;
meconium aspiration syndrome; congenital diaphragmatic hernia; lesion;
acute lung injury; cystic fibrosis; surgical wound; resection;
growth factor; smoke inhalation; HRG-alpha.
OS Homo sapiens.
XX US2001023241-A1.
XX
XX 20-SEP-2001.
XX
XX 02-FEB-2001; 2001US-0773517.
XX
XX 04-FEB-1998; 98US-073866P.
XX 02-FEB-1999; 99US-0243198.
XX
XX (SLIW/) SLIWOWSKI M X.
XX (KERN/) KERN J A.
XX
XX Sliwowski MX, Kern JA;
XX
XX WPI; 2001-595807/67.
XX N-PSDB; AAS18521.
XX
XX Inducing epithelial cell growth and/or proliferation, useful in the
treatment of respiratory disease, comprises use of heregulin ligand as
growth factor -
Claim 4; Fig 1; 68pp; English.
The invention relates to inducing epithelial cell growth and/or
proliferation comprising contacting a normal epithelial cell, which
expresses HER2, HER3 and/or HER4 (preferably HER2/HER3, HER2/HER4,
HER3/HER4, HER3 or HER4) receptors with an isolated ligand such as
heregulin (HRG) which activates HER2, HER3 and/or HER4 receptors. The
method is used for inducing epithelial cell (preferably lung cell) growth
and/or proliferation, for increasing lung surfactant protein A and for
treating respiratory distress or emphysema, for treating chronic
obstructive pulmonary disease (COPD) e.g. chronic bronchitis, asthma,
neonatal pulmonary diseases including neonatal respiratory distress
syndrome, meconium aspiration syndrome, chronic lung disease of the
neonate, congenital diaphragmatic hernia and acute lung injuries
including smoke or chemical inhalation, pneumonitis due to aspiration,
radiation, near drowning, cystic fibrosis and other epithelial cell
trauma diseases including injuries associated with surgical wounds and
resections, ulcers, lesions and tissue tears as normal epithelial cell
growth factors. For treating infants/neonates with respiratory distress
as well as youth and adult with poor lung function due to lung injury
or damage. The HRG binds with varying and very high affinity to the
HER2, HER3 and/or HER4 receptors. The method stimulates growth and
proliferation of the epithelial cells, repairing and re-establishing
the cellular barriers of organs and allowing the affected tissue to
develop normal physiological functions more quickly. Hence the method
improves oxygenation and speeds of the development of a barrier to
infection while treating lung cells that are damaged by inhalation of
smoke resulting in emphysema. It also facilitates regeneration of

CC epithelial cells. The present sequence represents the amino acid sequence
CC of human heregulin HRG-alpha.
XX
SQ Sequence 669 AA;
Query Match 100.0%; Score 3470; DB 22; Length 669;
Best Local Similarity 100.0%; Pred. No. 3.7e-215;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ARAPQGRSLSPSRDKLPNPPIRALGNPSPAPRAVRVRSVSGEMSKERKGRGKGGK 60
Db 1 ARAPQGRSLSPSRDKLPNPPIRALGNPSPAPRAVRVRSVSGEMSKERKGRGKGGK 60
Qy 61 ERSGKKPESAGSQSPALPRLKEMKQESAAGSKLVLCETSSYSSLRKFWKNGNE 120
Db 61 ERSGKKPESAGSQSPALPRLKEMKQESAAGSKLVLCETSSYSSLRKFWKNGNE 120
Qy 121 LNRNKKPQNIQKPKKSELINIKASLDGGEYCKVSKLGNDSANITIVESNEII 180
Db 121 LNRNKKPQNIQKPKKSELINIKASLDGGEYCKVSKLGNDSANITIVESNEII 180
Qy 181 TGMPESTEGAYVSSSPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
Db 181 TGMPESTEGAYVSSSPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
Qy 241 FMVKDLNPSRYLCKQPGFTGARTENPMKVQNEKAELYOKRVLTITIGICALLV 300
Db 241 FMVKDLNPSRYLCKQPGFTGARTENPMKVQNEKAELYOKRVLTITIGICALLV 300
Qy 301 GIMCVAYCKTKQKKLHDLRSLRSERNMMNIANGPHHPNPENVLQVNOYVSKN 360
Db 301 GIMCVAYCKTKQKKLHDLRSLRSERNMMNIANGPHHPNPENVLQVNOYVSKN 360
Qy 361 VISSEHIVERAETSFTSHYTSTAHSSTTTPSHSWNGHTESILSHSVIVMSSV 420
Db 361 VISSEHIVERAETSFTSHYTSTAHSSTTTPSHSWNGHTESILSHSVIVMSSV 420
Qy 421 ENSRHSSPTGGPRLNGTGPRCNSFLRHARETSDSYRSPHSERYVSAMTTPARMSP 480
Db 421 ENSRHSSPTGGPRLNGTGPRCNSFLRHARETSDSYRSPHSERYVSAMTTPARMSP 480
Qy 481 VDFHTSPSPKSP 540
Db 481 VDFHTSPSPKSP 540
Qy 541 QFSFHNPAHDSNLSPLRIVEDEEYETQYEPAQEPVKKLSRAKRTKPNNGHI 600
Db 541 QFSFHNPAHDSNLSPLRIVEDEEYETQYEPAQEPVKKLSRAKRTKPNNGHI 600
Qy 601 ANRLEVDNNTSSQSSNSESETEDEVEDPFLGIONPLAASLEATPAFLADSRTPAG 660
Db 601 ANRLEVDNNTSSQSSNSESETEDEVEDPFLGIONPLAASLEATPAFLADSRTPAG 660
Qy 661 RFTQEEIQ 669
Db 661 RFTQEEIQ 669
RESULT 3
AAU09886
ID AAU09886 standard; Protein; 669 AA.
XX
XX AAU09886;
XX
XX 20-APR-1993 (first entry)
XX Human heregulin-alpha deduced from clone lambda gt10 her16.
XX
XX p185-HER2 ligand; her2 proto-oncogene; breast cancer; EGF receptor;
KW epidermal growth factor; tyrosine kinase-like glycoprotein;
KW HRG-alpha.
XX
XX Homo sapiens.
OS

XX FH Key Location/Qualifiers
 FT Domain 288..309
 FT /label= "transmembrane"
 FT /note= "probable"
 FT Region 226..265
 FT /label= "EGF_motif"
 FT /note= "contains 6 cysteines"
 FT Modified-site 42..43
 FT /note= "serine-glycine dipeptide potential
 FT glycosaminoglycan addition site -
 FT not part of mature HRG-alpha sequence"
 FT Modified-site 64..65
 FT /note= "serine-glycine dipeptide potential
 FT glycosaminoglycan addition site"
 FT Modified-site 151..152
 FT /note= "serine-glycine dipeptide potential
 FT glycosaminoglycan addition site"
 FT Modified-site 164..166
 FT /note= "N-linked glycosylation site"
 FT Modified-site 170..172
 FT /note= "N-linked glycosylation site"
 FT Modified-site 208..210
 FT /note= "N-linked glycosylation site"
 FT Modified-site 437..439
 FT /note= "N-linked glycosylation site"
 FT Modified-site 609..611
 FT /note= "N-linked glycosylation site"
 FT Modified-site 209..221
 FT /note= "serine-threonine potential
 FT O-glycosylation sites"
 FT Region 1..44
 FT /note= "deduced from ORF but initiating MET is
 FT at position 45 and the processed
 FT N-terminal residue is Ser46"
 XX WO9220798-A.
 XX 26-NOV-1992.
 XX 21-MAY-1992; 92WO-US04295.
 XX 24-MAY-1991; 91US-0705256.
 XX 25-SEP-1991; 91US-0765212.
 XX 08-NOV-1991; 91US-0790801.
 XX 06-MAR-1992; 92US-0847743.
 XX 11-MAY-1992; 92US-0880917.
 XX (GENE-) GENETECH INC.
 PT Holmes WE, Vandlen RL;
 XX WPI: 1992-415776/50.
 XX P-PSDB; AAR29570.
 PT New polypeptide(s) heregulin(s) bind the p185HER2 receptor - for
 PT diagnosing and treating tumours and screening for agonists or
 PT antagonists for binding to p185HER2
 XX Example 3; Fig 4; 139pp; English.
 XX Heregulin-alpha was isolated from the supernatant of human breast
 CC carcinoma MDA-MB-231. After purification, a major peak of tyrosine
 CC phosphorylation activity was correlated with a 45kd protein. The
 CC protein was cleaved with Lysine-C and one of the resulting peptide
 CC fragments (C-15) was used to design an oligonucleotide probe (see
 CC AAQ31547). The probe was labelled and used to screen an oligo
 CC dr-primed cDNA library constructed from human MDA-MB-231 cell mRNA
 CC in lambda gt10. Two positive clones, designated lambda gt10 her16
 CC and lambda gt10 her13, were identified and found to be identical.
 CC The sequence of lambda gt10 her16 contains a single ORF of 669
 CC amino acids. No stop codon was found in the translated sequence.
 CC See also AAQ31542-Q31548.

XX SQ Sequence 669 AA;
 Query Match 99.9%; Score 3466; DB 13; Length 669;
 Best Local Similarity 99.9%; Pred. No. 6.6e-215;
 Matches 668; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ARAPQGRSLSPSRDKLPNPPIRALGPNSPAPRAVVRVSVSGEMSEKRGKGGKK 60
 DB 1 ARAPQGRSLSPSRDKLPNPPIRALGPNSPAPRAVVRVSVSGEMSEKRGKGGKK 60
 QY 61 ERSGKKPESAAAGSQSPALPPRLKEMKQESAAAGSKLVLRCTSEYSSLRKFKNGNE 120
 DB 61 ERSGKKPESAAAGSQSPALPPRLKEMKQESAAAGSKLVLRCTSEYSSLRKFKNGNE 120
 QY 121 LNRKNKPNQIKQKPGKSELINAKSLADSEYCKVSKLGNDSASANTIVESNEII 180
 DB 121 LNRKNKPNQIKQKPGKSELINAKSLADSEYCKVSKLGNDSASANTIVESNEII 180
 QY 181 TCMFASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLVKCAKEKTCVNGGEC 240
 DB 181 TCMFASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLVKCAKEKTCVNGGEC 240
 QY 241 FMVKDLSNPSRYLCKQPGFTGARTENVMKVQNOERAEELYQKRVLTITGICIALLV 300
 DB 241 FMVKDLSNPSRYLCKQPGFTGARTENVMKVQNOERAEELYQKRVLTITGICIALLV 300
 QY 301 GIMCVVAYCKTKQKRLHDLRQSLRSENNMMNIANGPHHPNPPENQVLYVSKN 360
 DB 301 GIMCVVAYCKTKQKRLHDLRQSLRSENNMMNIANGPHHPNPPENQVLYVSKN 360
 QY 361 VISSEHIVERAEATSFSTSHVTSTAHSTTQTPSHSWSNGHTSILSESHSVIMSSV 420
 DB 361 VISSEHIVERAEATSFSTSHVTSTAHSTTQTPSHSWSNGHTSILSESHSVIMSSV 420
 QY 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYRDSPHSERVYSAMTTPARMS 480
 DB 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYRDSPHSERVYSAMTTPARMS 480
 QY 481 VDFHTPSSPKPPSEMSPPVSSMTVSMPSMAVSPMEERPLLVTPLRLREKFDHHPQ 540
 DB 481 VDFHTPSSPKPPSEMSPPVSSMTVSMPSMAVSPMEERPLLVTPLRLREKFDHHPQ 540
 QY 541 QFSFPHNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANRRARTRPNNGHI 600
 DB 541 QFSFPHNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANRRARTRPNNGHI 600
 QY 601 ANRLEVDNTSSQSSNSETEDEVRGDTPLGTONPLAASLEATPAFRADSTNPAG 660
 DB 601 ANRLEVDNTSSQSSNSETEDEVRGDTPLGTONPLAASLEATPAFRADSTNPAG 660
 QY 661 RFTQEEIQ 669
 DB 661 RFTQEEIQ 669
 RESULT 4
 AAR29571
 ID AAR29571 standard; Protein; 675 AA.
 XX AC AAR29571;
 XX 20-APR-1993 (first entry)
 DT Human heregulin-beta-1 deduced from clone lambda her11.1db1.
 DE p185-HER2 ligand; her2 proto-oncogene; breast cancer; EGF receptor;
 XX epidermal growth factor; tyrosine kinase-like glycoprotein;
 KW HRG-beta1.
 XX Homo sapiens.
 XX Location/Qualifiers
 FH Key

Disclosure; Fig 2; 141pp; English.

The patent discloses a method for inducing hair cell generation, or inner-ear-supporting cell growth, regeneration, and/or proliferation, by heregulin (HRG), the ligand for HER2/HER3 dimeric receptors. Heregulin proteins function as activators of HER-2 oncogene and result from alternate splicing of a single gene mapped to chromosome 8p. The two major types, alpha and beta HRG's are based on two variant EGF-like (epidermal growth factor) domains, that differ in their C-terminal ends. This method can be used to increase the number of inner-ear-supporting cells and for treatment of hair cell related hearing disorders and disease states associated with tissue damage, e.g. ototoxic injury, acoustic assault, degenerative hearing loss, balance impairments, damage associated with surgery or physical injury and inner ear disorders related to hair cell dysfunction. The present sequence is the human heregulin, HRG-beta protein, derived from a genomic library. This sequence can be used to construct several variants.

Sequence 675 AA;

Query Match 94.6%; Score 3283.5; DB 21; Length 675;
Best Local Similarity 96.7%; Pred. No. 3.8e-203;
Matches 638; Conservative 2; Mismatches 15; Indels 5;

QY	15	DKLFPPNPIRALGPNSPAPRAV	RSVSGEMSERKGGKGGK	KKKRGSGKKPESAA	GS 74
D6	1	DKLFPPNPIRALGPNSPAPRAV	RSVSGEMSERKGGKGGK	KKKRGSGKKPESAA	GS 60

Qy	75	QSPALPRLKEMKQESAAGSKLVLCRTSSEYSSLRFKWFKNGNELNRKKNKPNIKIQK	134
Db	61	QSPALPQLKEMKQESAAGSKLVLCRTSSEYSSLRFKWFKNGNELNRKKNKPNIKIQK	120

[illegible]

QY	195	181	195
Db	ESPIRISVTEGANTSSSTSTTGTGTHLVKCAEKEKTCVNGGECFVVKDLSNP	ESPIRISVTEGANTSSSTSTTGTGTHLVKCAEKEKTCVNGGECFVVKDLSNP	ESPIRISVTEGANTSSSTSTTGTGTHLVKCAEKEKTCVNGGECFVVKDLSNP
	SYLC	SYLC	SYLC
	254	240	254

Qy	255	KCPGTTGARTENVPKVNQ-----EKAELYQKRVLTTITGICIALLVVGMCVAYC	309
		: :	
Db	241	KCPNETGDRCNQYNASFYKHLGIEMEAELYQKRVLTTITGICIALLVVGMCVAYC	300

QY	310	KT	KQ	BK	KL	HD	BL	DL	RL	QSL	RS	SR	NR	NN	MM	MM	T	ANG	PH	PH	PH	PP	PP	PP	PN	VQ	LV	QY	YS	KN	VS	SS	EH	VE	369
	301	KT	KQ	BK	KL	HD	BL	DL	RL	QSL	RS	SR	NR	NN	MM	MM	T	ANG	PH	PH	PH	PP	PP	PP	PN	VQ	LV	QY	YS	KN	VS	SS	EH	VE	360

[illegible]

Qy	430	GGPGR	LNGTGGPRECNSFLRHARETPDSYRSPHSERYVSAMTTPARMSPVDFHTPSSP	489
Dd	421	GGPGR	LNGTGGPRECNSFLRHARETPDSYRSPHSERYVSAMTTPARMSPVDFHTPSSP	480

Qy	490	KSPPEMSPPVSSMTYMP	SMAVSPFMEERLLLVTP	PRLRKKKFDHHPQDSSFFHN	P 54.9
Db	481	KSPPEMSPPVSSMTYMP	SMAVSPFMEERPLLVTP	PRLRKKKFDHHPQDSSFFHN	P 54.0

QY	550	AHDSN	PAS	PLR	IVE	DEE	YET	QYEP	QYEP	KVKK	LAN	SRRA	KRT	KPN	GHIAN	LR	VD	SN	609
	541	AHDSN <th>PAS</th> <th>PLR</th> <th>IVE</th> <th>DEE</th> <th>YET</th> <th>QYEP</th> <th>QYEP</th> <th>KVKK</th> <th>LAN</th> <th>SRRA</th> <th>KRT</th> <th>KPN</th> <th>GHIAN</th> <th>LR</th> <th>VD</th> <th>SN</th> <th>600</th>	PAS	PLR	IVE	DEE	YET	QYEP	QYEP	KVKK	LAN	SRRA	KRT	KPN	GHIAN	LR	VD	SN	600

Qy	610	TSSQSNSESETE	DERVGEDP	PLGIONPLA	ASLEATPA	FRUADSR	TNPAGR	STQEEIQ	669
Db	601	TSSQSNSESETE	DERVGEDP <td>PLGIONPLA <td>ASLEATPA <td>FRUADSR <td>TNPAGR <td>STQEEIQ</td> <td>660</td> </td></td></td></td>	PLGIONPLA <td>ASLEATPA <td>FRUADSR <td>TNPAGR <td>STQEEIQ</td> <td>660</td> </td></td></td>	ASLEATPA <td>FRUADSR <td>TNPAGR <td>STQEEIQ</td> <td>660</td> </td></td>	FRUADSR <td>TNPAGR <td>STQEEIQ</td> <td>660</td> </td>	TNPAGR <td>STQEEIQ</td> <td>660</td>	STQEEIQ	660

RESULT 7
AA71199

AT THE COURT OF APPEALS, CIVIL DIVISION, NO. 07-3 AA.

XX AC

AA	21-SEP-2000 (first entry)
DT	
XX	
DE	Human Heregulin HRG-beta1 protein variant-22.

Heregulin; HRG beta1; human; mutant; mutein; inner-ear-supporting cell;
 activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;
 hair cell related hearing disorder; ototoxic injury; tissue damage;
 acoustic assault; degenerative hearing loss; balance impairment;
 treatment; surgical injury; physical injury; inner ear disorder.

OS	Homo sapiens.
OS	Synthetic.

XX	Key	Location/Qualifiers
FH	Misc-difference 31	
FT		/label= Initiator_methionine
FT	Misc-difference 243	

FT
FT
FT
FT
FT
FT

AA
PN WO200027426-A1.

AA PD 18-MAY-2000.

AA
PF 28-OCT-1999; 99WO-US25744.

AA 07-NOV-1998; 98US-0107522.
PR

AA (GETH) GENENTECH INC.

AA
PI
Gao W;AA
DR WPI; 2000-376313/32.

Method for inducing hair cell generation and inner-ear-supporting cell growth regeneration and proliferation, useful for treating hearing disorders -

PS Disclosure; Page -: 141pp; English.

The patent discloses a method for inducing hair cell generation, or inner-ear-supporting cell growth, regeneration, and/or proliferation, by heregulin (HRG), the ligand for HER2/HER3 dimeric receptors. Heregulin proteins function as activators of HER-2 oncogene and result from alternate splicing of a single gene mapped to chromosome 8p. The two major types, alpha and beta HRG's are based on two variant EGF-like (epidermal growth factor) domains, that differ in their C-terminal ends. This method can be used to increase the number of inner-ear-supporting cells and for treatment of hair cell related hearing disorders and disease states associated with tissue damage, e.g. ototoxic injury, acoustic assault, degenerative hearing loss, balance impairments, damage associated with surgery or physical injury and inner ear disorders related to hair cell dysfunction. The present sequence is the human heregulin ligand, HRG-beta1 protein variant-22.

CC NOTE: THIS SEQUENCE IS NOT FOUND IN THE SPECIFICATION AND IS DERIVED
CC FROM THE HRG-BETAL PROTEIN SEQUENCE FOUND IN FIGURE 2 (AAV711172).

XX	Sequence	675 AA;
SQ		

Query Match	94.6%	Score 3283.5;	DB 21;	Length 675;
Best Local Similarity	96.7%	Pred. No. 3.8e-203;		
Matches 638: Conservative	2: Mismatches	15: Indels	5:	

QY 15 DKLFNPPIRAGPNSPAPRAVRVRSVSGEMSERKEGRCKGKGGKKERGSCKKPSAAGS 74

DD	1	DKLF	FNP	IKALG	NSFAP	RAVK	KVERS	VS	GEMSE	KKK	EGK	GRG	GRK	KKR	KKR	GGSG	KKK	PESA	AGS	8
QY	75	QSPAL	PPRL	KEMKS	QESA	AGSK	LV	RCET	SSEY	SSLR	FKW	FKNG	NEL	NR	KNK	PQNI	KIQK	134		

Db 541 AIDS/SLPASP/RIVEDEYETQYEPAPQVKKLANSRAKRTKPNGHIANRLEVDN 600
 QY 610 TSSQSSNSESETEDEVDGTPFLGIONPLAASLEATPAFLADSRTPNAGRFSTQEEIQ 669
 Db 601 TSSQSSNSESETEDEVDGTPFLGIONPLAASLEATPAFLADSRTPNAGRFSTQEEIQ 660

RESULT 9

AA71203
 ID AAY71203 standard; Protein; 675 AA.

AC AAY71203;

DT 21-SEP-2000 (first entry)

DE Human Heregulin HRG-beta1 protein variant-26.

EW Heregulin; HRG beta1; human; mutant; mutein; inner-ear-supporting cell;
 activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;
 hair cell related hearing disorder; ototoxic injury; tissue damage;
 acoustic assault; degenerative hearing loss; balance impairment;
 treatment; surgical injury; physical injury; inner ear disorder.

KW Homo sapiens.

OS Synthetic.

QS Key Location/Qualifiers

FT Misc-difference 31

FT /label= Initiator_methionine

FT Misc-difference 256

FT /note= "Wild type Met replaced with Ile;

FT This variable site corresponds to position 226, of the

FT HRG-beta1 mature protein, after initiation Met (M31)."

PN WO200027426-A1.

PD 18-MAY-2000.

XX 28-OCT-1999; 99WO-US25744.

XX 07-NOV-1998; 98US-0107522.

XX (GETH) GENENTECH INC.

PI Gao W;

DR WPI; 2000-376313/32.

Method for inducing hair cell generation and inner-ear-supporting cell
 growth regeneration and proliferation, useful for treating hearing
 disorders

Disclosure; Page -; 141pp; English.

The patent discloses a method for inducing hair cell generation, or
 inner-ear-supporting cell growth, regeneration, and/or proliferation, by
 heregulin (HRG), the ligand for HER2/HER3 dimeric receptors. Heregulin
 proteins function as activators of HER-2 oncogene and result from
 alternate splicing of a single gene mapped to chromosome 8p. The two
 major types, alpha and beta HRG's are based on two variant EGF-like
 (epidermal growth factor) domains, that differ in their C-terminal ends.
 This method can be used to increase the number of inner-ear-supporting
 cells and for treatment of hair cell related hearing disorders and
 disease states associated with tissue damage, e.g. ototoxic injury,
 acoustic assault, degenerative hearing loss, balance impairments, damage
 associated with surgery or physical injury and inner ear disorders
 related to hair cell dysfunction. The present sequence is the human
 heregulin ligand, HRG-beta1 protein variant-26.
 Note: This sequence is not found in the specification and is derived
 from the HRG-beta1 protein sequence found in Figure 2 (AA711172).

XX Sequence 675 AA;

Query Match 94.6%; Score 3283.5; DB 21; Length 675;
 Best Local Similarity 96.7%; Pred. NO. 3.8e-203;
 Matches 639; Conservative 4; Mismatches 11; Indels 7; Gaps 2;
 QY 15 DKLPNPTRALGPNSPAPRAVYRVSVEGSEKRGKGKGGKKGKGSKKKESAAAGS 74
 Db 1 DKLPNPTRALGPNSPAPRAVYRVSVEGSEKRGKGKGGKKGKGSKKKESAAAGS 60
 QY 75 QSPALPPRLKEMKQESAAAGSKLVLRCTSEYSSSLRKFKNGLNKRKNKPNQIKQ 134
 Db 61 QSPALPPQLKEMKQESAAAGSKLVLRCTSEYSSSLRKFKNGLNKRKNKPNQIKQ 120
 QY 135 KPGKSELINKASLADSGEYMKVSKLGNDSASANTIVESNIIITGMPASTEGAYVSS 194
 Db 121 KPGKSELINKASLADSGEYMKVSKLGNDSASANTIVESNIIITGMPASTEGAYVSS 180
 QY 195 ESPRISVSTEGANTSSSTSTTGTSHLVKCAKEKTCFVNGGECFVMDLSPSYLC 254
 Db 181 ESPRISVSTEGANTSSSTSTTGTSHLVKCAKEKTCFVNGGECFVMDLSPSYLC 240
 QY 255 KCQPGFTGARTENV-----PMKVQNOEKAEEYQKRVLTITGICIAALLVVGIMCVAY 308
 Db 241 KCPNEFTGDRQNTVYIASFYKHLGIEFME-AEELYQKRVLTITGICIAALLVVGIMCVAY 299
 QY 309 CKTKKQKKLHDLRQLRSERNMMNTANGPHHPNPPENQVQVYVSKNVISSSEHIV 368
 Db 300 CKTKKQKKLHDLRQLRSERNMMNTANGPHHPNPPENQVQVYVSKNVISSSEHIV 359
 QY 369 EREAETSFSTSHYTSTAHHSTVTQTPSHSNNGHTESSILSESHSVIVMSSVENSRRSSP 428
 Db 360 EREAETSFSTSHYTSTAHHSTVTQTPSHSNNGHTESSILSESHSVIVMSSVENSRRSSP 419
 QY 429 TGGPRGLNGTGGPRECNFLRHARETSDYRSPHSERYYSAMTTPARMSPVDFHTPSS 488
 Db 420 TGGPRGLNGTGGPRECNFLRHARETSDYRSPHSERYYSAMTTPARMSPVDFHTPSS 479
 QY 489 PKSPPEMSPPVSSMTVSMPSMAVSPPMEEERPLLVTTPRLREKKKDDHHPQOQSSHHN 548
 Db 480 PKSPPEMSPPVSSMTVSMPSMAVSPPMEEERPLLVTTPRLREKKKDDHHPQOQSSHHN 539
 QY 549 PAHDSNLPASPLRIVEDEEYETQYEPAPQVKKLANSRAKRTKPNGHIANRLEVDS 608
 Db 540 PAHDSNLPASPLRIVEDEEYETQYEPAPQVKKLANSRAKRTKPNGHIANRLEVDS 599
 QY 609 NTSSQSSNSESETEDEVDGTPFLGIONPLAASLEATPAFLADSRTPNAGRFSTQEEI 668
 Db 600 NTSSQSSNSESETEDEVDGTPFLGIONPLAASLEATPAFLADSRTPNAGRFSTQEEI 659
 QY 669 Q 669
 Db 660 Q 660

RESULT 10

AAU09887
 ID AAU09887 standard; Protein; 675 AA.

XX AAU09887;

DT 26-FEB-2002 (first entry)

XX Human heregulin, HRG-beta1.

DE Human heregulin, HRG-beta1.

XX Human; heregulin; antiasthmatic; antiinflammatory; vulnery; antiulcer;
 epithelial cell growth; HER2; HER3; HRG; lung cell; asthma;
 lung surfactant protein A; respiratory distress; emphysema; pneumonitis;
 chronic obstructive pulmonary disease; COPD; chronic bronchitis; ulcer;
 neonatal pulmonary disease; neonatal respiratory distress syndrome;
 meconium aspiration syndrome; congenital diaphragmatic hernia; lesion;
 acute lung injury; cystic fibrosis; surgical wound; resection;
 growth factor; smoke inhalation; HRG-beta1.

XX

OS Homo sapiens.
PN US2001023241-A1.
XX 20-SEP-2001.
XX 02-FEB-2001; 2001US-07733517.
XX 04-FEB-1998; 98US-073866P.
PR 02-FEB-1999; 99US-0243198.
XX (SLIW/) SLINKOWSKI M X.
PA (KERN/) KERN J A.
XX
XX Sliwowski MX, Kern JA;
XX WPI; 2001-595807/67.
DR N-PSDB; AAS18522.
XX
Inducing epithelial cell growth and/or proliferation, useful in the treatment of respiratory disease, comprises use of heregulin ligand as growth factor -
Claim 4; Fig 2; 68pp; English.
The invention relates to inducing epithelial cell growth and/or proliferation comprising contacting a normal epithelial cell, which expresses HER2, HER3 and/or HER4 (preferably HER2/HER3, HER2/HER4, HER3/HER4, or HER4) receptors with an isolated ligand such as heregulin (HRG) which activates HER2, HER3 and/or HER4 receptors. The method is used for inducing epithelial cell (preferably lung cell) growth and/or proliferation, for increasing lung surfactant protein A and for treating respiratory distress or emphysema, for treating chronic obstructive pulmonary disease (COPD) e.g. chronic bronchitis, asthma, neonatal pulmonary diseases including neonatal respiratory distress syndrome, meconium aspiration syndrome, chronic lung disease of the neonate, congenital diaphragmatic hernia and acute lung injuries including smoke or chemical inhalation, pneumonitis due to aspiration, radiation, near drowning, cystic fibrosis and other epithelial cell trauma diseases including injuries associated with surgical wounds and resections, ulcers, lesions and tissue tears as normal epithelial cell growth factors. For treating infants/neonates with respiratory distress as well as youth and adult with poor lung function due to lung injury or damage. The HRG binds with varying and very high affinity to the HER2, HER3 and/or HER4 receptors. The method stimulates growth and proliferation of the epithelial cells, repairing and re-establishing the cellular barriers of organs and allowing the affected tissue to develop normal physiological functions more quickly. Hence the method improves oxygenation and speeds of the development of a barrier to infection while treating lung cells that are damaged by inhalation of smoke resulting in emphysema. It also facilitates regeneration of epithelial cells. The present sequence represents the amino acid sequence of human heregulin HRG-beta1.
SQ Sequence 675 AA;
Query Match 94.6%; Score 3283.5; DB 22; Length 675;
Best Local Similarity 96.7%; Pred. No. 3.8e-203;
Matches 638; Conservative 2; Mismatches 15; Indels 5; Gaps 1;
QY 15 DKLPNPPIRALGPNPAPRAVVERSVGEMERKGRGKGGKRGSGKKPESAGS 74
DB 1 DKLPNPPIRALGPNPAPRAVVERSVGEMERKGRGKGGKRGSGKKPESAGS 60
QY 75 QSPALPRLKEMKQESAGSKLVLCETSSSYSLRFKFWKNGELNKNKPNQIKIQK 134
DB 61 QSPALPQLKEMKQESAGSKLVLCETSSSYSLRFKFWKNGELNKNKPNQIKIQK 120
QY 135 KPGKSELRLINKASLADSGYMKVTSKLGNDGSANITIVESNEIITGMPASTEGAYVSS 194
DB 121 KPGKSELRLINKASLADSGYMKVTSKLGNDGSANITIVESNEIITGMPASTEGAYVSS 180
QY 195 ESPIRISVTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGECFMVKDLSNPSRYLC 254

DB 181 ESPIRISVTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGECFMVKDLSNPSRYLC 240
QY 255 KCPGFTGARTENVPKVNQ-----EKAEELYQKRVLTITGICITALLVVGIMCVAYC 309
DB 241 KCPNEFTGDRCONVYMASFYKHLGIEFMEAEELYQKRVLTITGICITALLVVGIMCVAYC 300
QY 310 KTKKQKRLHDLRLQSLRSERNNMNIANGPHHPNPPENQVLYVNVYVSKNVISSEHIVE 369
DB 301 KTKKQKRLHDLRLQSLRSERNNMNIANGPHHPNPPENQVLYVNVYVSKNVISSEHIVE 360
QY 370 REAETSFTSHYTSTAHTSTTVTQTPSHWSNGHTESILSESHSVIVMSVENSRSPT 429
DB 361 REAETSFTSHYTSTAHTSTTVTQTPSHWSNGHTESILSESHSVIVMSVENSRSPT 420
QY 430 GPGRLNGTGGPRECNFLRHARETSDYRSPHSEYVSAMTTPARMSPVDFHTPSSP 489
DB 421 GPGRLNGTGGPRECNFLRHARETSDYRSPHSEYVSAMTTPARMSPVDFHTPSSP 480
QY 490 KSPPEMSPPVSSMTVSMFMAVSPFMEERPLLLVTPPRLREKFKDHPHPQPFSSFHNP 549
DB 481 KSPPEMSPPVSSMTVSMFMAVSPFMEERPLLLVTPPRLREKFKDHPHPQPFSSFHNP 540
QY 550 AHDNSLPASPLRIVEDEEYETQYEPAQEPVKKLANSSRAKRTKPNGHIANRLEVDN 609
DB 541 AHDNSLPASPLRIVEDEEYETQYEPAQEPVKKLANSSRAKRTKPNGHIANRLEVDN 600
QY 610 TSSQSSNSETEDEVRGDDTFLGIONPLAASLEATPAFLADSTNPAGRFSTQEEIQ 669
DB 601 TSSQSSNSETEDEVRGDDTFLGIONPLAASLEATPAFLADSTNPAGRFSTQEEIQ 660
RESULT 11
AAY71202
ID AAY71202 standard; Protein; 675 AA.
XX AAY71202;
AC AAY71202;
DT 21-SEP-2000 (first entry)
XX Human Heregulin HRG-beta1 protein variant-25.
DE Heregulin; HRG beta1; human; mutant; muten; inner-ear-supporting cell;
KW activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;
KW hair cell related hearing disorder; ototoxic injury; tissue damage;
KW acoustic assault; degenerative hearing loss; balance impairment;
KW treatment; surgical injury; physical injury; inner ear disorder.
XX Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 31
FT FT /label= Initiator_methionine
FT FT /label= His, Trp
FT FT /note= *this variable site corresponds to position 223 of
the HRG-beta1 mature protein, after initiator Met (M31).
PN WO200027426-A1.
XX 18-MAY-2000.
XX 28-OCT-1999; 99WO-US25744.
PF 07-NOV-1998; 98US-0107522.
XX (GETH) GENENTECH INC.
XX Gao W;
XX WPI; 2000-376313/32.
XX

Qy	135	KPGISELRINKASLADSGEYCKVIKSLGNDSSANITIVESNEIITGMPASTEGAYVSS	194
Db	121	KPGISELRINKASLADSGEYCKVIKSLGNDSSANITIVESNEIITGMPASTEGAYVSS	180
Qy	195	ESPIRISVSTEGANTSSSTSTSTGTSHLVCAEKEKTCVNGGBCFMVKDLSNPSRYLC	254
Db	181	ESPIRISVSTEGANTSSSTSTSTGTSHLVCAEKEKTCVNGGBCFMVKDLSNPSRYLC	240
Qy	255	KCPGFTGARCENY-----PMKVQNOKEABELYQKRLVLTITGICIAALLVVVGIMCVVAY	308
Db	241	KCPNEFTGDRQCMVTVIASFYKHLGIEFME-AEELYQKRLVLTITGICIAALLVVVGIMCVVAY	299
Qy	309	CKTKQRKKLHDLRQLRSRERNNMNANGPHHPNPPENVQLVQTVKKNVISSEHTV	368
Db	300	CKTKQRKKLHDLRQLRSRERNNMNANGPHHPNPPENVQLVQTVKKNVISSEHTV	359
Qy	369	EREAEISFSTSHYSTAHSTTVTQTPSHSNNGHTSEILSESHSVIVMSSVENSRRHSSP	428
Db	360	EREAEISFSTSHYSTAHSTTVTQTPSHSNNGHTSEILSESHSVIVMSSVENSRRHSSP	419
Qy	429	TGGPRGLNGTGGPRECNFSLRHARETDPDSYRDSPHSERYYSAMTTPARMSVPDPTTPSS	488
Db	420	TGGPRGLNGTGGPRECNFSLRHARETDPDSYRDSPHSERYYSAMTTPARMSVPDPTTPSS	479
Qy	489	PKSPPEMSPPVSSMTVSMPSMAYSPFMEERPLLVTTPRLREKKFDHHPQOFSFHHN	548
Db	480	PKSPPEMSPPVSSMTVSMPSMAYSPFMEERPLLVTTPRLREKKFDHHPQOFSFHHN	539
Qy	549	PAHDSNLSPLASPLRIVEDEEYETQVEYPAQEPVKKLSANRRAKTKPNGHIANLEVDS	608
Db	540	PAHDSNLSPLASPLRIVEDEEYETQVEYPAQEPVKKLSANRRAKTKPNGHIANLEVDS	599
Qy	609	NTSSQSNSESETERVGEDTTPFLGIONPLAASLEATPAFLADSRTPNAGRFSTQBEI	668
Db	600	NTSSQSNSESETERVGEDTTPFLGIONPLAASLEATPAFLADSRTPNAGRFSTQBEI	659
Qy	669	Q 669	
Db	660	Q 660	
RESULT 13			
ID	AA711188	standard; Protein; 675 AA.	
XX	AC		
XX	AA711188;		
XX	21-SEP-2000	(first entry)	
DE	Human Heregulin HRG-beta1 protein variant-11.		
DE	Heregulin; HRG beta1; human; mutant; mutein; Inner-ear-supporting cell;		
KW	activator; HER2 receptor; HER3 receptor; HER-2 oncogene; Chromosome 8p;		
KW	hair cell related hearing disorder; ototoxic injury; tissue damage;		
KW	acoustic assault; degenerative hearing loss; balance impairment;		
KW	treatment; surgical injury; physical injury; inner ear disorder.		
XX			
OS	Homo sapiens.		
OS	Synthetic.		
XX			
Key	Location/Qualifiers		
FT	Misc-difference 31	/label- Initiator_methionine	
FT	Misc-difference 225	/note= "Wild type Glu replaced with Gln;	
FT		This variable site corresponds to position 195, of the	
FT		HRG-beta1 mature protein, after initiator Met (M31)"	
XX			
PN	WO200027426-A1.		
XX			
PD	18-MAY-2000.		
XX			

PF	28-OCT-1999;	99WO-US25744.	
XX			
PR	07-NOV-1998;	98US-0107522.	
XX			
PA	(GETH) GENENTECH INC.		
XX			
XX	Gao W;		
PI			
XX	WPI; 2000-376313/32.		
DR			
XX			
PT	Method for inducing hair cell generation and inner-ear-supporting cell		
PT	growth regeneration and proliferation, useful for treating hearing		
PT	disorders		
XX			
PS	Disclosure; Page -; 141pp; English.		
XX			
CC	The patent discloses a method for inducing hair cell generation, or		
CC	inner-ear-supporting cell growth, regeneration, and/or proliferation, by		
CC	heregulin (HRG), the ligand for HER2/HER3 dimeric receptors. Heregulin		
CC	proteins function as activators of HER-2 oncogene and result from		
CC	alternate splicing of a single gene mapped to chromosome 8p. The two		
CC	major types, alpha and beta HRG's are based on two variant EGF-like		
CC	(epidermal growth factor) domains, that differ in their C-terminal ends.		
CC	This method can be used to increase the number of inner-ear-supporting		
CC	cells and for treatment of hair cell related hearing disorders and		
CC	disease states associated with tissue damage, e.g. ototoxic injury,		
CC	acoustic assault, degenerative hearing loss, balance impairments, damage		
CC	associated with surgery or physical injury and inner ear disorders		
CC	related to hair cell dysfunction. The present sequence is the human		
CC	heregulin ligand, HRG-beta1 protein variant-11.		
CC	Note: This sequence is not found in the specification and is derived		
CC	from the HRG-beta1 protein sequence found in Figure 2 (AA71172).		
XX			
XX	Sequence 675 AA;		
QY	Query Match	94.5%; Score 3280.5; DB 21; Length 675;	
DB	Best Local Similarity	96.5%; Pred. No. 6e-203;	
DB	Matches 637; Conservative	3; Mismatches 15; Indels 5; Gaps 1;	
QY	15	DKLFPNPIRALGPSNAPRAVRSVSGEMSRKEGRGKGKKGKRGSGKKPESAAGS	74
DB	1	DKLFPNPIRALGPSNAPRAVRSVSGEMSRKEGRGKGKKGKRGSGKKPESAAGS	60
QY	75	QSPALPRLKEMKSOESAAGSKLVLRCTSESYSLSLRFKFKNGELNLRNKKPNQNIQK	134
DB	61	QSPALPQLKEMKSOESAAGSKLVLRCTSESYSLSLRFKFKNGELNLRNKKPNQNIQK	120
QY	135	KPKSELRLINKASLADSGEYCKVIKSLGNDNSANITIVESNEITGMPASTEGAYVSS	194
DB	121	KPKSELRLINKASLADSGEYCKVIKSLGNDNSANITIVESNEITGMPASTEGAYVSS	180
QY	195	ESPIRISVSTEGANTSSSTSTGTGTHLVKCAEKETFCVNGGECFMVKDLNPSRYLC	254
DB	181	ESPIRISVSTEGANTSSSTSTGTGTHLVKCAEKETFCVNGGECFMVKDLNPSRYLC	240
QY	255	KCQPGTGTGARTENVPKVNQ-----EKAEEYLQKRVLITGICIALLVVGMVAVC	309
DB	241	KCPNEFTGDCQNYVMASFYKHLGIEFMEAEELYQKRVLTITGICIALLVVGMVAVC	300
QY	310	KTKQKRLKDLRLQSLRSENNMMNIANGPHNPENPNVQVNVYVSKNVTSSSEHIVE	369
DB	301	KTKQKRLKDLRLQSLRSENNMMNIANGPHNPENPNVQVNVYVSKNVTSSSEHIVE	360
QY	370	REAETSFSTSHYTSTAHHSTTVTQTPSHWSNGHTESILSESHSVIVMSVENSRRSSPT	429
DB	361	REAETSFSTSHYTSTAHHSTTVTQTPSHWSNGHTESILSESHSVIVMSVENSRRSSPT	420
QY	430	GGPRGLNGTGGPRECNSTFLRHARETSDYRDSPHSEYRYVSAMTTPARMSPVDFHTPSSP	489
DB	421	GGPRGLNGTGGPRECNSTFLRHARETSDYRDSPHSEYRYVSAMTTPARMSPVDFHTPSSP	480
QY	490	KSPPEMSPPVSWMTVSMPSMAYSPFMEERPLLLVTPPLRLREKKTDPHPOQSSSHHPN	549

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:44:11 ; Search time 27 Seconds
(without alignments)
729.035 Million cell updates/sec

Title: US-10-022-609-11

Perfect score: 3470

Sequence: 1 ARAPQRSLSPSRDKLPN.....RLADRTNPAGRFSTQIEIQ 669

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA: *
- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
 - 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
 - 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
 - 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
 - 5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
 - 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3470	100.0	669	1	US-07-847-743B-8
2	3470	100.0	669	1	US-07-847-743B-13
3	3470	100.0	669	1	US-08-456-201-8
4	3470	100.0	669	1	US-08-456-201-13
5	3470	100.0	669	2	US-08-456-201-11
6	3470	100.0	669	2	US-08-330-161-11
7	3470	100.0	669	2	US-08-456-241-8
8	3470	100.0	669	2	US-08-456-241-13
9	3470	100.0	669	2	US-08-440-401-11
10	3470	100.0	669	2	US-08-419-878B-11
11	3470	100.0	669	4	US-09-173-480-11
12	3470	100.0	669	5	PCT-US92-04295A-8
13	3283.5	94.6	732	1	US-07-847-743B-9
14	3283.5	94.6	732	1	US-07-847-743B-9
15	3283.5	94.6	732	2	US-08-456-201-9
16	3283.5	94.6	732	2	US-08-456-241-9
17	3247	93.6	625	5	PCT-US92-04295A-9
18	3247	93.6	625	1	US-07-847-743B-26
19	3247	93.6	625	1	US-08-456-201-26
20	3247	93.6	625	5	US-08-456-241-26
21	3247	93.6	640	4	US-08-467-602-256
22	3232.5	93.2	649	4	US-08-467-602-267
23	3227	93.0	660	4	US-08-467-602-264
24	3222.5	92.9	669	4	US-08-467-602-273
25	3213.5	92.6	687	4	US-08-467-602-274
26	3199	92.2	696	4	US-08-467-602-258
27	3193.5	92.0	707	4	US-08-467-602-262

28	3189	91.9	716	4	US-08-467-602-274	Sequence 274, App
29	3149	90.7	646	4	US-08-467-602-270	Sequence 270, App
30	3141.5	90.5	637	4	US-08-467-602-258	Sequence 258, App
31	3133.5	90.3	645	3	US-08-753-007A-10	Sequence 10, Appl
32	3133.5	90.3	645	4	US-09-398-496-10	Sequence 10, Appl
33	3131.5	90.2	637	1	US-07-847-743B-28	Sequence 28, Appl
34	3131.5	90.2	637	1	US-08-456-201-28	Sequence 28, Appl
35	3131.5	90.2	637	2	US-08-456-241-28	Sequence 28, Appl
36	3131.5	90.2	637	5	PCT-US92-04295A-28	Sequence 28, Appl
37	3129.5	90.2	645	1	US-07-847-743B-27	Sequence 27, Appl
38	3129.5	90.2	645	1	US-08-456-201-27	Sequence 27, Appl
39	3129.5	90.2	645	1	US-08-428-926-4	Sequence 4, Appl
40	3129.5	90.2	645	1	US-08-428-927-4	Sequence 4, Appl
41	3129.5	90.2	645	1	US-08-428-298-4	Sequence 4, Appl
42	3129.5	90.2	645	1	US-08-339-517-4	Sequence 4, Appl
43	3129.5	90.2	645	2	US-08-456-241-27	Sequence 27, Appl
44	3129.5	90.2	645	4	US-09-020-880-93	Sequence 93, Appl
45	3129.5	90.2	645	4	US-09-101-544-93	Sequence 93, Appl

ALIGNMENTS

RESULT 1
US-07-847-743B-8
; Sequence 8, Application US/07847743B
; Patent No. 5367060
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/847.743B
; FILING DATE: 19920306
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; FILING DATE: 24-May-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/765212
; FILING DATE: 25-Sep-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790801
; FILING DATE: 08-No. 5367060-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hensley, Max D.
; REGISTRATION NUMBER: 27,043
; REFERENCE/DOCKET NUMBER: 712P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1489
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear

Query Match 100.0%; Score 3470; DB 1; Length 669;

Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGGK 60
DB 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGGK 60

QY 61 ERSGKKPESAAGSQSPALPPRLKEMKQESAAGSKLVLCETSSYSSILRFKFWKNGNE 120
DB 61 ERSGKKPESAAGSQSPALPPRLKEMKQESAAGSKLVLCETSSYSSILRFKFWKNGNE 120

QY 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMKVSKLGNDSASANITIVESNEII 180
DB 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMKVSKLGNDSASANITIVESNEII 180

QY 181 TGMPTAGYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TGMPTAGYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240

QY 241 FVVKDLSNPSRYLCKQPGTGARTENPMKVQNOEKABELYQKRVLTITGICIALVY 300
DB 241 FVVKDLSNPSRYLCKQPGTGARTENPMKVQNOEKABELYQKRVLTITGICIALVY 300

QY 301 GIMCVVAYCKTKQKPKSELRLINKASLADSGEYMKVSKLGNDSASANITIVESNEII 360
DB 301 GIMCVVAYCKTKQKPKSELRLINKASLADSGEYMKVSKLGNDSASANITIVESNEII 360

QY 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHWSNGHTESSILSESHSVIYVMSV 420
DB 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHWSNGHTESSILSESHSVIYVMSV 420

QY 421 ENSRHSSPTGGPRGLNGTGGPRECSFLRHARETDPDYSRDSPHSRYRYSAMTTPARMSP 480
DB 421 ENSRHSSPTGGPRGLNGTGGPRECSFLRHARETDPDYSRDSPHSRYRYSAMTTPARMSP 480

QY 481 VDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLREKKFDHHPQ 540
DB 481 VDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLREKKFDHHPQ 540

QY 541 QFSSFHNPADHSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANSSRAKRTKPNNGHI 600
DB 541 QFSSFHNPADHSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANSSRAKRTKPNNGHI 600

QY 601 ANRLEVDNSTSSQSSNSETEDEYVGEDTPTFLGIONPLAASLEATPAFLADSRTPNAG 660
DB 601 ANRLEVDNSTSSQSSNSETEDEYVGEDTPTFLGIONPLAASLEATPAFLADSRTPNAG 660

QY 661 RFSTQREIQ 669
DB 661 RFSTQREIQ 669

RESULT 2
US-07-847-743B-13
Sequence 13, Application US/07847743B
Patent No. 5367060

GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Structure, Production and Use of
TITLE OF INVENTION: Heregulin
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/847,743B
FILING DATE: 19920306
CLASSIFICATION: 424
PRIOR APPLICATION DATA: 07/705256
FILING DATE: 24-May-1991
APPLICATION NUMBER: 07/765212
FILING DATE: 25-Sep-1991
PRIOR APPLICATION DATA: 07/790801
APPLICATION NUMBER: 08-No. 5367060-1991
FILING DATE: 08-No. 5367060-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,043
REFERENCE/DOCKET NUMBER: 712P3
TELEPHONE: 415/266-1489
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-847-743B-13

Query Match 100.0%; Score 3470; DB 1; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGGK 60
DB 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGGK 60

QY 61 ERSGKKPESAAGSQSPALPPRLKEMKQESAAGSKLVLCETSSYSSILRFKFWKNGNE 120
DB 61 ERSGKKPESAAGSQSPALPPRLKEMKQESAAGSKLVLCETSSYSSILRFKFWKNGNE 120

QY 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMKVSKLGNDSASANITIVESNEII 180
DB 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMKVSKLGNDSASANITIVESNEII 180

QY 181 TGMPTAGYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TGMPTAGYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240

QY 241 FVVKDLSNPSRYLCKQPGTGARTENPMKVQNOEKABELYQKRVLTITGICIALVY 300
DB 241 FVVKDLSNPSRYLCKQPGTGARTENPMKVQNOEKABELYQKRVLTITGICIALVY 300

QY 301 GIMCVVAYCKTKQKPKSELRLINKASLADSGEYMKVSKLGNDSASANITIVESNEII 360
DB 301 GIMCVVAYCKTKQKPKSELRLINKASLADSGEYMKVSKLGNDSASANITIVESNEII 360

QY 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHWSNGHTESSILSESHSVIYVMSV 420
DB 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHWSNGHTESSILSESHSVIYVMSV 420

QY 421 ENSRHSSPTGGPRGLNGTGGPRECSFLRHARETDPDYSRDSPHSRYRYSAMTTPARMSP 480
DB 421 ENSRHSSPTGGPRGLNGTGGPRECSFLRHARETDPDYSRDSPHSRYRYSAMTTPARMSP 480

QY 481 VDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLREKKFDHHPQ 540
DB 481 VDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLREKKFDHHPQ 540

QY 541 QFSSFHNPADHSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANSSRAKRTKPNNGHI 600
DB 541 QFSSFHNPADHSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANSSRAKRTKPNNGHI 600

QY 601 ANRLEVDNSTSSQSSNSETEDEYVGEDTPTFLGIONPLAASLEATPAFLADSRTPNAG 660
DB 601 ANRLEVDNSTSSQSSNSETEDEYVGEDTPTFLGIONPLAASLEATPAFLADSRTPNAG 660

Db 601 ANRLEVDNTSSQSSSETEDEVRGDTFFLQIONPLAASLEATPAFLADSRTPAG 660
QY 661 RFSTQEEIQ 669
Db 661 RFSTQEEIQ 669

RESULT 3
US-08-456-201-8
; Sequence 8, Application US/08456201
; Patent No. 5641869
; GENERAL INFORMATION:
; APPLICANT: Vandlen, Richard L.,
; APPLICANT: Holmes, William E.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin 2 Ligands
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,201
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,145
; FILING DATE:
; PRIOR APPLICATION NUMBER: 07/880,917
; APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; APPLICATION DATA:
; APPLICATION NUMBER: 07/765212
; APPLICATION DATA:
; APPLICATION NUMBER: 07/790801
; ATTORNEY/AGENT INFORMATION:
; NAME: Hensley, Max D.
; REGISTRATION NUMBER: 27,034
; REFERENCE/DOCKET NUMBER: 712P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1489
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-456-201-8

Query Match 100.0%; Score 3470; DB 1: Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAFORGRSLSPSRDKLPNPNIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGGK 60
Db 1 ARAFORGRSLSPSRDKLPNPNIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGGK 60
QY 61 ERGSKKPESAGSOSPAPRLPKEMKSOESAAAGSKLVRCETSEYSSSLRPFKFKNGNE 120
Db 61 ERGSKKPESAGSOSPAPRLPKEMKSOESAAAGSKLVRCETSEYSSSLRPFKFKNGNE 120
QY 121 LNRNKNPONIQRKPKGSELRINKASLADSGEYWKVSKLGNDSASANTIVNESNII 180

Db 121 LNRNKNPONIQRKPKGSELRINKASLADSGEYWKVSKLGNDSASANTIVNESNII 180
QY 181 TCMPEASTGAYVSSSEPIRISVSTEGANTSSSTSTGTSHLVKCAEKEKTFVNGGEC 240
Db 181 TCMPEASTGAYVSSSEPIRISVSTEGANTSSSTSTGTSHLVKCAEKEKTFVNGGEC 240
QY 241 FMVKDLNPSRYLCKCQPGFTGARCETENVMKVONQEKAEELYQKRVLTITGICIALLV 300
Db 241 FMVKDLNPSRYLCKCQPGFTGARCETENVMKVONQEKAEELYQKRVLTITGICIALLV 300
QY 301 GIMCVVAYCKTKKORRKLHDLRSLRSERNMNIANGPHHPNPPENVOLVNOYVSKN 360
Db 301 GIMCVVAYCKTKKORRKLHDLRSLRSERNMNIANGPHHPNPPENVOLVNOYVSKN 360
QY 361 VISSEHIVEREAETSFSTSHYTSTAHHSTTTVQTPSHSWNSNGHTESILSESHSVIVMSSV 420
Db 361 VISSEHIVEREAETSFSTSHYTSTAHHSTTTVQTPSHSWNSNGHTESILSESHSVIVMSSV 420
QY 421 ENSRHSSPTGGPRGLNCTGGPRECNSFLRHARETPODSYRSPHSERYVSAMTTPARMSP 480
Db 421 ENSRHSSPTGGPRGLNCTGGPRECNSFLRHARETPODSYRSPHSERYVSAMTTPARMSP 480
QY 481 VDFHTPPSPKPPSPSPVSSMTVSMPSMAVSPFMEERERPLLVTPLRREKKFDDHPQ 540
Db 481 VDFHTPPSPKPPSPSPVSSMTVSMPSMAVSPFMEERERPLLVTPLRREKKFDDHPQ 540
QY 541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQETPEYPAQEPVKKLSRRRAKRTKPNCHI 600
Db 541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQETPEYPAQEPVKKLSRRRAKRTKPNCHI 600
QY 601 ANRLEVDNTSSQSSSETEDEVRGDTFFLQIONPLAASLEATPAFLADSRTPAG 660
Db 601 ANRLEVDNTSSQSSSETEDEVRGDTFFLQIONPLAASLEATPAFLADSRTPAG 660
QY 661 RFSTQEEIQ 669
Db 661 RFSTQEEIQ 669

RESULT 4
US-08-456-201-13
; Sequence 13, Application US/08456201
; Patent No. 5641869
; GENERAL INFORMATION:
; APPLICANT: Vandlen, Richard L.,
; APPLICANT: Holmes, William E.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin 2 Ligands
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,201
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,145
; FILING DATE:
; PRIOR APPLICATION NUMBER: 07/880,917
; APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; APPLICATION DATA:
; APPLICATION NUMBER: 07/765212
; APPLICATION DATA:
; APPLICATION NUMBER: 07/790801
; APPLICATION DATA:
; APPLICATION NUMBER: 07/765212

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/790801
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hensley, Max D.
;; REGISTRATION NUMBER: 27,034
;; REFERENCE/DOCKET NUMBER: 712P4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/266-1489
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 669 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; US-08-456-201-13

Query Match 100.0%; Score 3470; DB 1; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ARAPQGRSLSPSRDLFPNIPRALGPNSPAPRAVRSVSGEMSERKEGRGKGKK 60
1 ARAPQGRSLSPSRDLFPNIPRALGPNSPAPRAVRSVSGEMSERKEGRGKGKK 60
QY 61 ERSGKKPESAGSQSPALPPRLKEMKQESAAGSKLVLCETSSYSSLRKFWKNGNE 120
Db 61 ERSGKKPESAGSQSPALPPRLKEMKQESAAGSKLVLCETSSYSSLRKFWKNGNE 120
QY 121 LNRKNKPNQIKOKPKGKSELRLNKASLADSGEYMCVISKLGNDASANITIVESNEII 180
Db 121 LNRKNKPNQIKOKPKGKSELRLNKASLADSGEYMCVISKLGNDASANITIVESNEII 180
QY 181 TGMPTAGAYVSSSPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
Db 181 TGMPTAGAYVSSSPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FMVKDLSNPSRYLCKQCPGTGARTENPMKVQNOEKAELYQKRVLTITGICIALVY 300
Db 241 FMVKDLSNPSRYLCKQCPGTGARTENPMKVQNOEKAELYQKRVLTITGICIALVY 300
QY 301 GIMCVVAYCKTKQKRLKHLDRQLRSRERNMMNTANGPHHPNPPENVQVYVSKN 360
Db 301 GIMCVVAYCKTKQKRLKHLDRQLRSRERNMMNTANGPHHPNPPENVQVYVSKN 360
QY 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHSWSNGHTESILSESHSVIYVSSV 420
Db 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHSWSNGHTESILSESHSVIYVSSV 420
QY 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYEDSPHSERYYSAMTTPARMSP 480
Db 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYEDSPHSERYYSAMTTPARMSP 480
QY 481 VDPHTPSSPKPPSPSEMSPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLEKKFDHPQ 540
Db 481 VDPHTPSSPKPPSPSEMSPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLEKKFDHPQ 540
QY 541 QFSFHHNPAHDSNLSLPASPLRIVEDEYETQYEPAPQEPVKKLANSRAKRTKPNGHI 600
Db 541 QFSFHHNPAHDSNLSLPASPLRIVEDEYETQYEPAPQEPVKKLANSRAKRTKPNGHI 600
QY 601 ANRLEYDSTSSQSSNSESETEDEYVGTDPFLGIONPLAASLEATPAFLADSRTPNAG 660
Db 601 ANRLEYDSTSSQSSNSESETEDEYVGTDPFLGIONPLAASLEATPAFLADSRTPNAG 660
QY 661 RFTQEQEIQ 669
Db 661 RFTQEQEIQ 669

RESULT 5
US-08-330-161-11
; Sequence 11, Application US/08330161

;; Patent No. 5834229
;; GENERAL INFORMATION:
;; APPLICANT: Vandlen, Richard
;; APPLICANT: Holmes, William
;; TITLE OF INVENTION: Structure, Production and Use of Heregulin 2 Ligands
;; NUMBER OF SEQUENCES: 17
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 460 Point San Bruno Blvd
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: patin (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/330,161
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/035430
;; FILING DATE: 22-MAR-1993
;; APPLICATION NUMBER: 07/705256
;; FILING DATE: 24-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Adler, Carolyn R.
;; REGISTRATION NUMBER: 32,324
;; REFERENCE/DOCKET NUMBER: 712C1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/225-2614
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 669 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; US-08-330-161-11

Query Match 100.0%; Score 3470; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDLFPNIPRALGPNSPAPRAVRSVSGEMSERKEGRGKGKK 60
Db 1 ARAPQGRSLSPSRDLFPNIPRALGPNSPAPRAVRSVSGEMSERKEGRGKGKK 60
QY 61 ERSGKKPESAGSQSPALPPRLKEMKQESAAGSKLVLCETSSYSSLRKFWKNGNE 120
Db 61 ERSGKKPESAGSQSPALPPRLKEMKQESAAGSKLVLCETSSYSSLRKFWKNGNE 120
QY 121 LNRKNKPNQIKOKPKGKSELRLNKASLADSGEYMCVISKLGNDASANITIVESNEII 180
Db 121 LNRKNKPNQIKOKPKGKSELRLNKASLADSGEYMCVISKLGNDASANITIVESNEII 180
QY 181 TGMPTAGAYVSSSPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
Db 181 TGMPTAGAYVSSSPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FMVKDLSNPSRYLCKQCPGTGARTENPMKVQNOEKAELYQKRVLTITGICIALVY 300
Db 241 FMVKDLSNPSRYLCKQCPGTGARTENPMKVQNOEKAELYQKRVLTITGICIALVY 300
QY 301 GIMCVVAYCKTKQKRLKHLDRQLRSRERNMMNTANGPHHPNPPENVQVYVSKN 360
Db 301 GIMCVVAYCKTKQKRLKHLDRQLRSRERNMMNTANGPHHPNPPENVQVYVSKN 360
QY 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHSWSNGHTESILSESHSVIYVSSV 420
Db 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHSWSNGHTESILSESHSVIYVSSV 420


```
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,241
FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/126145
FILING DATE: 23-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/880917
FILING DATE: 11-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847743
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/790801
FILING DATE: 08-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/765212
FILING DATE: 25-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705256
FILING DATE: 24-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 712P4CID1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-9894
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-456-241-13

Query Match 100.0%; Score 3470; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVVERSVSGEMSERKEGRKGKGGK 60
DB 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVVERSVSGEMSERKEGRKGKGGK 60
QY 61 ERGSGKKPSAAGSGSPALPRLKEMKSOESAAGSKVLVRCSTSEYSSLRKFWKNGNE 120
DB 61 ERGSGKKPSAAGSGSPALPRLKEMKSOESAAGSKVLVRCSTSEYSSLRKFWKNGNE 120
QY 121 LARKNKPQIKQKPKGKSELINKASLADSGEYKVKLSKLGNDASANITIVESNETI 180
DB 121 LARKNKPQIKQKPKGKSELINKASLADSGEYKVKLSKLGNDASANITIVESNETI 180
QY 181 TGMPTASTEGAYSSPISRYSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCFVNGGEC 240
DB 181 TGMPTASTEGAYSSPISRYSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCFVNGGEC 240
QY 241 FVVKDLSNPSRYLCKQCPGFTGARTENYPMKVQKQKAEELYQKRVLTITGICIALVY 300
DB 241 FVVKDLSNPSRYLCKQCPGFTGARTENYPMKVQKQKAEELYQKRVLTITGICIALVY 300
QY 301 GIMCVVAYCKTKQKRLKDLRLQSLRSERNNMNIANGPHHPNPPENVLVNOYVSKN 360
DB 301 GIMCVVAYCKTKQKRLKDLRLQSLRSERNNMNIANGPHHPNPPENVLVNOYVSKN 360

301 GIMCVVAYCKTKQKRLKDLRLQSLRSERNNMNIANGPHHPNPPENVLVNOYVSKN 360
361 VISSSEHIVERAEATSFSTSHYTSTAHHSTTVQTSPSHWSNGHSTESILSESHSVIYVSSV 420
361 VISSSEHIVERAEATSFSTSHYTSTAHHSTTVQTSPSHWSNGHSTESILSESHSVIYVSSV 420
421 ENSRHSSPTGGPRGLNGTGGPRECNSFLRHARETTPDSYRDSPHSERYYSAMTTPARMSP 480
421 ENSRHSSPTGGPRGLNGTGGPRECNSFLRHARETTPDSYRDSPHSERYYSAMTTPARMSP 480
481 VDFHTPSSPKSPSEMSPVSVMTVSMYSPMAYSPFMEERPLLLVTPPLRREKKKFDHHPQ 540
481 VDFHTPSSPKSPSEMSPVSVMTVSMYSPMAYSPFMEERPLLLVTPPLRREKKKFDHHPQ 540
541 QFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETQVEPAQEPVKKLANSSRAKRTKPNNGHI 600
541 QFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETQVEPAQEPVKKLANSSRAKRTKPNNGHI 600
601 ANRLEVDNTSSQSSNSESETEDEVRGDETPFLGIONPLAASLEATPAFLADSRTPNAG 660
601 ANRLEVDNTSSQSSNSESETEDEVRGDETPFLGIONPLAASLEATPAFLADSRTPNAG 660
661 RESTQEEIQ 669
661 RESTQEEIQ 669

RESULT 8
US-08-440-401-11
Sequence 11, Application US/08440401
Patent No. 5856110
GENERAL INFORMATION:
APPLICANT: Vandlen, Richard L.
APPLICANT: Holmes, William E.
TITLE OF INVENTION: Structure, Production and Use of
TITLE OF INVENTION: Heregulin 2 ligands
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,401
FILING DATE: 12-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/330161
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/035430
FILING DATE: 22-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705256
FILING DATE: 24-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 712C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-9994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 amino acids
```

TYPE: amino acid
TOPOLOGY: linear
US-08-440-401-11

Query Match
Best Local Similarity 100.0%; Score 3470; DB 2; Length 669;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGKKK 60
DB 1 ARAPQGRSLSPSRDKLPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGKKK 60
QY 61 ERSGKKPESAGSOSPALPRLKEMKSOEASGKLVLCETSEYSSSLRFKFKNGNE 120
DB 61 ERSGKKPESAGSOSPALPRLKEMKSOEASGKLVLCETSEYSSSLRFKFKNGNE 120
QY 121 LNRNKPONIKQKPKGKSELRLKASLADSGEYCKVSKLGNDSASANTIVESNEII 180
DB 121 LNRNKPONIKQKPKGKSELRLKASLADSGEYCKVSKLGNDSASANTIVESNEII 180
QY 181 TGMPTAGYVSSPESPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TGMPTAGYVSSPESPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FMVDLSNPSRYLCKCPGFTGARTENVPKVNQOEKAEELYOKRVLTITGICALLV 300
DB 241 FMVDLSNPSRYLCKCPGFTGARTENVPKVNQOEKAEELYOKRVLTITGICALLV 300
QY 301 GIMCVAYCKTKKQKLLHDLRLQSLRSERNMNIANGPHHPNPPENVOLVQYVSKN 360
DB 301 GIMCVAYCKTKKQKLLHDLRLQSLRSERNMNIANGPHHPNPPENVOLVQYVSKN 360
QY 361 VISSEHIVERAEETSFSTSHYTSHTAHSTTTTQPSHWSNGHTESILSESHVIMSSV 420
DB 361 VISSEHIVERAEETSFSTSHYTSHTAHSTTTTQPSHWSNGHTESILSESHVIMSSV 420
QY 421 ENSRHSSPTGPRGLNGTGPRCNSFLRHARETPOSYRDSPHSERVVSAMTTPARMSP 480
DB 421 ENSRHSSPTGPRGLNGTGPRCNSFLRHARETPOSYRDSPHSERVVSAMTTPARMSP 480
QY 481 VDFHTPSSPKPPSEMSPVSMVSMPSMAVSPFMEERPLLVTTPRLREKFFDHPQ 540
DB 481 VDFHTPSSPKPPSEMSPVSMVSMPSMAVSPFMEERPLLVTTPRLREKFFDHPQ 540
QY 541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANRRKRTKPNNGHI 600
DB 541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANRRKRTKPNNGHI 600
QY 601 ANRLEVDNTSSQSSNESETEDETVGDTPLGIONPLAASLEATPAFLADSTNPAG 660
DB 601 ANRLEVDNTSSQSSNESETEDETVGDTPLGIONPLAASLEATPAFLADSTNPAG 660
QY 661 RFSQEEIQ 669
DB 661 RFSQEEIQ 669

RESULT 9
US-08-419-878B-11
Sequence 11 Application US/08419878B
Patent No 5859206
GENERAL INFORMATION:
APPLICANT: Vandlen, Richard L.
TITLE OF INVENTION: Antibodies Specific For Heregulin 2-alpha
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/419,878B
FILING DATE: 11-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/330161
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/035430
FILING DATE: 22-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705256
FILING DATE: 24-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0712C2D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-419-878B-11

Query Match 100.0%; Score 3470; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGKKK 60
DB 1 ARAPQGRSLSPSRDKLPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGKKK 60
QY 61 ERSGKKPESAGSOSPALPRLKEMKSOEASGKLVLCETSEYSSSLRFKFKNGNE 120
DB 61 ERSGKKPESAGSOSPALPRLKEMKSOEASGKLVLCETSEYSSSLRFKFKNGNE 120
QY 121 LNRNKPONIKQKPKGKSELRLKASLADSGEYCKVSKLGNDSASANTIVESNEII 180
DB 121 LNRNKPONIKQKPKGKSELRLKASLADSGEYCKVSKLGNDSASANTIVESNEII 180
QY 181 TGMPTAGYVSSPESPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TGMPTAGYVSSPESPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FMVDLSNPSRYLCKCPGFTGARTENVPKVNQOEKAEELYOKRVLTITGICALLV 300
DB 241 FMVDLSNPSRYLCKCPGFTGARTENVPKVNQOEKAEELYOKRVLTITGICALLV 300
QY 301 GIMCVAYCKTKKQKLLHDLRLQSLRSERNMNIANGPHHPNPPENVOLVQYVSKN 360
DB 301 GIMCVAYCKTKKQKLLHDLRLQSLRSERNMNIANGPHHPNPPENVOLVQYVSKN 360
QY 361 VISSEHIVERAEETSFSTSHYTSHTAHSTTTTQPSHWSNGHTESILSESHVIMSSV 420
DB 361 VISSEHIVERAEETSFSTSHYTSHTAHSTTTTQPSHWSNGHTESILSESHVIMSSV 420
QY 421 ENSRHSSPTGPRGLNGTGPRCNSFLRHARETPOSYRDSPHSERVVSAMTTPARMSP 480
DB 421 ENSRHSSPTGPRGLNGTGPRCNSFLRHARETPOSYRDSPHSERVVSAMTTPARMSP 480
QY 481 VDFHTPSSPKPPSEMSPVSMVSMPSMAVSPFMEERPLLVTTPRLREKFFDHPQ 540
DB 481 VDFHTPSSPKPPSEMSPVSMVSMPSMAVSPFMEERPLLVTTPRLREKFFDHPQ 540
QY 541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANRRKRTKPNNGHI 600

Db 541 QFSFHNPAHDSNLSPLRIVEDEYETQYEPAPVKKLANSRAKTKPNHGI 600
|||||

Qy 601 ANRLEVDNTSSOSSNSESETEDEYVGEDEPFLGIONPLAASLEATPAFLADSRNPAG 660
|||||

Db 601 ANRLEVDNTSSOSSNSESETEDEYVGEDEPFLGIONPLAASLEATPAFLADSRNPAG 660
|||||

Qy 661 RFSTQEEIQ 669
|||||

Db 661 RFSTQEEIQ 669
|||||

RESULT 10

US-09-173-480-11

; Sequence 11, Application US/09173480

; Patent No. 6399746

; GENERAL INFORMATION:

; APPLICANT: Vandlen, Richard L.

; TITLE OF INVENTION: Structure, Production and Use of

; Heregulin 2 Ligands

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/173,480

; FILING DATE: 14-Oct-1998

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/440,401

; FILING DATE: <Unknown>

; APPLICATION NUMBER: 08/035430

; FILING DATE: 22-MAR-1993

; APPLICATION NUMBER: 07/705256

; FILING DATE: 24-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Lee, Wendy M.

; REGISTRATION NUMBER: 00,000

; REFERENCE/DOCKET NUMBER: 712C3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/225-1994

; TELEFAX: 415/952-9881

; TELE: 910/371-7168

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 669 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-09-173-480-11

Query Match 100.0%; Score 3470; DB 4; Length 669;

Best Local Similarity 100.0%; Pred. NO. 1.4e-245;

Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARAPOGRSLSPSRDKLFNPTRALGNPSPAPRAVVERSVSGEMSERKEGKGGKK 60

|||||

Db 1 ARAPOGRSLSPSRDKLFNPTRALGNPSPAPRAVVERSVSGEMSERKEGKGGKK 60

|||||

Qy 61 ERSGKPKPSAAGSQSPALPPRLKEMKQSAAGSKVLVRCETSSYSLSLRFKFKNGNE 120

|||||

Db 61 ERSGKPKPSAAGSQSPALPPRLKEMKQSAAGSKVLVRCETSSYSLSLRFKFKNGNE 120

|||||

Qy 121 LNRKNRPONIKIQKPKSELINKASLADSGEYMKVISKLGNDASANITIVESNEII 180
|||||

Db 121 LNRKNRPONIKIQKPKSELINKASLADSGEYMKVISKLGNDASANITIVESNEII 180
|||||

Qy 181 TGMPTAGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAEKTEFCVNGGEC 240
|||||

Db 181 TGMPTAGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAEKTEFCVNGGEC 240
|||||

Qy 241 FMVKDLNPSRYLCKQPGFTGARCENPMKVONQEKAEELYQKRVLTITGICIALLVV 300
|||||

Db 241 FMVKDLNPSRYLCKQPGFTGARCENPMKVONQEKAEELYQKRVLTITGICIALLVV 300
|||||

Qy 301 GIMCVVAYCKTKQRKKLHDLRQLRSERNNMNIANGPHHPNPPENVQVYNQVSKN 360
|||||

Db 301 GIMCVVAYCKTKQRKKLHDLRQLRSERNNMNIANGPHHPNPPENVQVYNQVSKN 360
|||||

Qy 361 VISSEHIVERAEATSFSTSHYHSTTAHSTTVTQTPSHSWSNGHTESILSESHSVIYVSSV 420
|||||

Db 361 VISSEHIVERAEATSFSTSHYHSTTAHSTTVTQTPSHSWSNGHTESILSESHSVIYVSSV 420
|||||

Qy 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYRDSPHSERYVSAMTTPARMSP 480
|||||

Db 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYRDSPHSERYVSAMTTPARMSP 480
|||||

Qy 481 VDFHTSPSPKPPSEMSPPVSSMTVSMVSPMAYSPFMEERPLLLVTPPRLREKKFDDHHPQ 540
|||||

Db 481 VDFHTSPSPKPPSEMSPPVSSMTVSMVSPMAYSPFMEERPLLLVTPPRLREKKFDDHHPQ 540
|||||

Qy 541 QFSFHNPAHDSNLSPLRIVEDEYETQYEPAPVKKLANSRAKTKPNHGI 600
|||||

Db 541 QFSFHNPAHDSNLSPLRIVEDEYETQYEPAPVKKLANSRAKTKPNHGI 600
|||||

Qy 601 ANRLEVDNTSSOSSNSESETEDEYVGEDEPFLGIONPLAASLEATPAFLADSRNPAG 660
|||||

Db 601 ANRLEVDNTSSOSSNSESETEDEYVGEDEPFLGIONPLAASLEATPAFLADSRNPAG 660
|||||

Qy 661 RFSTQEEIQ 669
|||||

Db 661 RFSTQEEIQ 669
|||||

RESULT 11

PCT-US92-04295A-8

; Sequence 8, Application PC/TUS9204295A

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; TITLE OF INVENTION: Structure, Production and Use of

; HEREGULIN

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/04295A

; FILING DATE: 19920521

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/705256

; FILING DATE: 24-MAY-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/765212

; FILING DATE: 25-SEP-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/790801

```

; FILING DATE: 08-NOV-1991
; PRIOR APPLICATION DATA: 07/847743
; APPLICATION NUMBER: 07/847743
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hensley, Max D.
; REGISTRATION NUMBER: 27,043
; REFERENCE/DOCKET NUMBER: 712P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US92-04295A-8

```

Query Match	100.0%;	Score 3470;	DB 5;	Length 669;
Best Local Similarity	100.0%;	Pred. No. 1.4e-245;		
Matches 669;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ARAPQGRSLSPSRDKLFPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGGKKK	60	
Db	1	ARAPQGRSLSPSRDKLFPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGGKKK	60	
QY	61	ERGSGBKPSAAGSQSPALPPRLKEMKQSQAAGSKVLVLCETSSSEYSSSLRFFKFKNGNE	120	
Db	61	ERGSGBKPSAAGSQSPALPPRLKEMKQSQAAGSKVLVLCETSSSEYSSSLRFFKFKNGNE	120	
QY	121	LNKNKPQNTKIQQKPKGSELINKASLADSGEYMKVISKLGNDASANITIVESNEII	180	
Db	121	LNKNKPQNTKIQQKPKGSELINKASLADSGEYMKVISKLGNDASANITIVESNEII	180	
QY	181	TGMPASTEGAYVSESPIRLSVSTEGANTSSSTSTSTGTSHLYKCAEKETFCVNGGEC	240	
Db	181	TGMPASTEGAYVSESPIRLSVSTEGANTSSSTSTSTGTSHLYKCAEKETFCVNGGEC	240	
QY	241	FMVKDLSNPARYLCKQPGFTGARCTENVPKMQNQEKAELYQKRVLTITGICIALLV	300	
Db	241	FMVKDLSNPARYLCKQPGFTGARCTENVPKMQNQEKAELYQKRVLTITGICIALLV	300	
QY	301	GIMCVVAYCKTKQKKLHDRLQSLRSERNMMNIANGPHHPNPPENVOLVNOQYVSKN	360	
Db	301	GIMCVVAYCKTKQKKLHDRLQSLRSERNMMNIANGPHHPNPPENVOLVNOQYVSKN	360	
QY	361	VISSEHIVERAEATSFSTSHYTSTAHHSTVTQTPSHWSNGHTESILSESHSVIVMSSV	420	
QY	361	VISSEHIVERAEATSFSTSHYTSTAHHSTVTQTPSHWSNGHTESILSESHSVIVMSSV	420	
QY	421	ENSRHSSPTGGPRGLNGTGGPRECNLSFLRHARETPOSYRDSPHSERVYSAMTTPARMSP	480	
Db	421	ENSRHSSPTGGPRGLNGTGGPRECNLSFLRHARETPOSYRDSPHSERVYSAMTTPARMSP	480	
QY	481	VDFHTPSSPKSPSPSEMSPPVSSMTVMSPMASVSPFMEERPLLLVTPPRLREKFKDHHQP	540	
Db	481	VDFHTPSSPKSPSPSEMSPPVSSMTVMSPMASVSPFMEERPLLLVTPPRLREKFKDHHQP	540	
QY	541	QFSSFFHNPAHDSNSLPASPLRIVEDEEYETTQYEPAQEPVKKLANRRAKRTPKNGHI	600	
Db	541	QFSSFFHNPAHDSNSLPASPLRIVEDEEYETTQYEPAQEPVKKLANRRAKRTPKNGHI	600	
QY	601	ANRLEVDSNTSSQSSNSESETEDERVEDTFFGICQNPLAASLEATPAFLRADSRTPNAG	660	
Db	601	ANRLEVDSNTSSQSSNSESETEDERVEDTFFGICQNPLAASLEATPAFLRADSRTPNAG	660	
QY	661	RFTSQBEIQ 669		
Db	661	RFTSQBEIQ 669		

RESULT 12

```

PCT-US92-04295A-13
: Sequence 13, Application PC/TUS9204295A
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc.
: TITLE OF INVENTION: Structure, Production and Use of
: TITLE OF INVENTION: Heregulin
: NUMBER OF SEQUENCES: 30
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/04295A
: FILING DATE: 19920521
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/705256
: FILING DATE: 24-MAY-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/765212
: FILING DATE: 25-SEP-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/790801
: FILING DATE: 08-NOV-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/847743
: FILING DATE: 06-MAR-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Hensley, Max D.
: REGISTRATION NUMBER: 27,043
: REFERENCE/DOCKET NUMBER: 712P4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 669 amino acids
: TYPE: AMINO ACIDS
: TOPOLOGY: linear
: PCT-US92-04295A-13

```

[illegible]

Db 1 DKLFNPIRALGPNPAPRAVVERSVSGEMSERGKGGKKGKSGKKPESAAGS 60
QY 75 QSPALPPRLKEMKQESAAGSKLVLRCTSSYSSLRPKFWKNGNLRKKNKPNQIKIOK 134
Db 61 QSPALPPQLKEMKQESAAGSKLVLRCTSSYSSLRPKFWKNGNLRKKNKPNQIKIOK 120
QY 135 KPGKSELIRINKASLADSGEYMKVISKLGNDASANITIVESNEIITGMPASTEGAYVSS 194
Db 121 KPGKSELIRINKASLADSGEYMKVISKLGNDASANITIVESNEIITGMPASTEGAYVSS 180
QY 195 ESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGECFVWKDLSNPSRYLC 254
Db 181 ESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGECFVWKDLSNPSRYLC 240
QY 255 KCQPGFTGARCENYPMKVQNQ-----EKAELYOKRVLTITGICIALLVGIMCVVAYC 309
Db 241 KCPNEFTGDRQNYWAFYKHLGTFEAEELYOKRVLTITGICIALLVGIMCVVAYC 300
310 KTKKQKKLHDLRQSLRSERNMMNIANGPHHPNPPPPENVOLVNOYVSKNVISSEHIVE 369
301 KTKKQKKLHDLRQSLRSERNMMNIANGPHHPNPPPPENVOLVNOYVSKNVISSEHIVE 360
370 REAETSFSSTSHYTSTAHHSTTVTQPSHSWNGHTESILSESHSVIVMSSVENSRSPT 429
Db 361 REAETSFSSTSHYTSTAHHSTTVTQPSHSWNGHTESILSESHSVIVMSSVENSRSPT 420
430 GGPRGLNGTGGPRCNSFLRHARETPDSYRDSPHSERVVSAMTTPARMSPVDFHTPSSP 489
Db 421 GGPRGLNGTGGPRCNSFLRHARETPDSYRDSPHSERVVSAMTTPARMSPVDFHTPSSP 480
QY 490 KSPPEMSPVSSMTVSMPSMAVSPFMEERPLLVTPPRLREKFDHHPQOQSSFHHP 549
Db 481 KSPPEMSPVSSMTVSMPSMAVSPFMEERPLLVTPPRLREKFDHHPQOQSSFHHP 540
QY 550 AHDSNLPASPLRIVEDEYEYEQEPAQEPVKKLANRRAKRTPNGHIANRLEVDN 609
Db 541 AHDSNLPASPLRIVEDEYEYEQEPAQEPVKKLANRRAKRTPNGHIANRLEVDN 600
QY 610 TSSQSSNSESETERVGEDTPFLGIONPLAASLEATPAFLADSRTPNAGRSTQEEIQ 669
Db 601 TSSQSSNSESETERVGEDTPFLGIONPLAASLEATPAFLADSRTPNAGRSTQEEIQ 660

Search completed: July 9, 2003, 10:48:41
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:45:26 ; Search time 28 Seconds
(without alignments)
2781.601 Million cell updates/sec

Title: US-10-022-609-11
Perfect score: 3470
Sequence: 1 ARAPQRSLSPSRDLFPN.....RLADSRNPAGRSTQEEIQ 669

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/RCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3470	100.0	669	9	US-10-022-609-11
2	3470	100.0	669	10	US-09-773-517-1
3	3470	100.0	669	10	US-09-792-025-1
4	3470	100.0	669	10	US-09-849-868-1
5	3283.5	94.6	675	10	US-09-773-517-3
6	3283.5	94.6	675	10	US-09-792-025-3
7	3283.5	94.6	675	10	US-09-849-868-3
8	3153	90.9	632	9	US-09-946-807-16
9	3153	90.9	632	10	US-09-795-668-16
10	3153	90.9	632	10	US-09-795-668-16
11	3133.5	90.3	645	12	US-10-096-241-10
12	3131.5	90.2	637	10	US-09-773-517-5
13	3131.5	90.2	637	10	US-09-792-025-5
14	3131.5	90.2	637	10	US-09-849-868-5
15	3047.5	87.8	629	9	US-09-946-807-14
16	3047.5	87.8	629	10	US-09-795-668-14
17	3047.5	87.8	629	10	US-09-795-668-14
18	3039.5	87.6	637	9	US-09-946-807-13
19	3039.5	87.6	637	10	US-09-795-668-13

20 3039.5 87.6 637 10 US-09-795-668-13
21 2599.5 74.9 548 9 US-09-946-807-29
22 2599.5 74.9 548 10 US-09-795-668-29
23 2599.5 74.9 548 10 US-09-795-668-29
24 2150 62.0 456 9 US-09-946-807-17
25 2150 62.0 456 10 US-09-795-668-17
26 2150 62.0 456 10 US-09-795-668-17
27 2092.5 60.3 420 10 US-09-773-517-9
28 2092.5 60.3 420 10 US-09-792-025-9
29 2092.5 60.3 420 10 US-09-849-868-9
30 2057.5 59.3 422 12 US-10-096-241-9
31 1733 49.9 344 9 US-09-946-807-38
32 1733 49.9 344 10 US-09-795-668-38
33 1733 49.9 344 10 US-09-795-668-38
34 1161.5 33.5 252 9 US-09-946-807-28
35 1161.5 33.5 252 10 US-09-795-668-28
36 1161.5 33.5 252 10 US-09-795-668-28
37 1105 31.8 241 10 US-09-773-517-7
38 1105 31.8 241 10 US-09-792-025-7
39 1105 31.8 241 10 US-09-849-868-7
40 1090 31.4 239 9 US-09-946-807-18
41 1090 31.4 239 10 US-09-795-668-18
42 1090 31.4 239 10 US-09-795-668-18
43 1056 30.4 209 9 US-09-946-807-15
44 1056 30.4 209 10 US-09-795-668-15
45 1056 30.4 209 10 US-09-795-668-15

ALIGNMENTS

RESULT 1
US-10-022-609-11
; Sequence 11, Application US/10022609
; Publication No. US20030023035A1
; GENERAL INFORMATION:
; APPLICANT: Vandlen, Richard L.
; TITLE OF INVENTION: Structure, Production and Use of Heregulin 2 Ligands
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/022.609
; FILING DATE: 17-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/440.401
; FILING DATE: 12-MAY-1995
; APPLICATION NUMBER: 08/330161
; FILING DATE: 25-OCT-1994
; APPLICATION NUMBER: 08/035430
; FILING DATE: 22-MAR-1993
; APPLICATION NUMBER: 07/705256
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 712C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881

Sequence 13, Appl
Sequence 29, Appl
Sequence 29, Appl
Sequence 29, Appl
Sequence 17, Appl
Sequence 17, Appl
Sequence 17, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 38, Appl
Sequence 38, Appl
Sequence 28, Appl
Sequence 28, Appl
Sequence 28, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 15, Appl

TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-022-609-11

```
Query Match      100.0%; Score 3470; DB 9; Length 669;
Best Local Similarity 100.0%; Pred. No. 4e-173;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLFPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGKGGK 60
DB 1 ARAPQGRSLSPSRDKLFPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGKGGK 60
QY 61 ERSGKKPESAGSOSPALPPRLKEMKSOESAGSKLVLCRTSSEYSSLRKFWKNGNE 120
DB 61 ERSGKKPESAGSOSPALPPRLKEMKSOESAGSKLVLCRTSSEYSSLRKFWKNGNE 120
QY 121 LNRKNKPQIKIQKPKGKSELINRKASLADSGEYCKVSKLGNDSASANITIVESNEII 180
DB 121 LNRKNKPQIKIQKPKGKSELINRKASLADSGEYCKVSKLGNDSASANITIVESNEII 180
QY 181 TCMFASTEGAYVSSSPIRISVSTEGANTSSSTSTSTTCTSHLVKCAEKEKTCVNGGEC 240
DB 181 TCMFASTEGAYVSSSPIRISVSTEGANTSSSTSTSTTCTSHLVKCAEKEKTCVNGGEC 240
QY 241 FVVKDLSNPSRYLCKCPGFTGARTENVPKQVQKAEELYQKRVLTITGICITALLVY 300
DB 241 FVVKDLSNPSRYLCKCPGFTGARTENVPKQVQKAEELYQKRVLTITGICITALLVY 300
QY 301 GIMCVVAYCKTKQKRLKDLRLQSLRSERNNMNIANGPHHPNPPENVLVNOYVSKN 360
DB 301 GIMCVVAYCKTKQKRLKDLRLQSLRSERNNMNIANGPHHPNPPENVLVNOYVSKN 360
QY 361 VISSHIVEREAETSFSTSHYTSTAHHSTVTQTPSHWSNGHTESSILSESHSVIWMSSV 420
DB 361 VISSHIVEREAETSFSTSHYTSTAHHSTVTQTPSHWSNGHTESSILSESHSVIWMSSV 420
QY 421 ENSRHSSPTGGPRGLNGTGGPRECNSFLRHARETSDSYRDSPHSERYYSAMTTARMSP 480
DB 421 ENSRHSSPTGGPRGLNGTGGPRECNSFLRHARETSDSYRDSPHSERYYSAMTTARMSP 480
QY 481 VDFHTPSSPKSPPEMSPVSMTPSMAYSPFMEERPLLLVTPPLREKKKFDHHPQ 540
DB 481 VDFHTPSSPKSPPEMSPVSMTPSMAYSPFMEERPLLLVTPPLREKKKFDHHPQ 540
QY 541 QFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANSRRARTKPNNGHI 600
DB 541 QFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANSRRARTKPNNGHI 600
QY 601 ANRLEVDNSTSQSSNSESETEDEVRGDTPLGIONPLAASLEATPAFLADSTRTNAG 660
DB 601 ANRLEVDNSTSQSSNSESETEDEVRGDTPLGIONPLAASLEATPAFLADSTRTNAG 660
QY 661 RFTSQEEIQ 669
DB 661 RFTSQEEIQ 669
```

RESULT 2

US-09-773-517-1
; Sequence 1, Application US/09773517
; Patent No. US20010023241A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/09/773,517
; CURRENT FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/243,198
; PRIOR FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-773-517-1

```
Query Match      100.0%; Score 3470; DB 10; Length 669;
Best Local Similarity 100.0%; Pred. No. 4e-173;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLFPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGKGGK 60
DB 1 ARAPQGRSLSPSRDKLFPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGKGGK 60
QY 61 ERSGKKPESAGSOSPALPPRLKEMKSOESAGSKLVLCRTSSEYSSLRKFWKNGNE 120
DB 61 ERSGKKPESAGSOSPALPPRLKEMKSOESAGSKLVLCRTSSEYSSLRKFWKNGNE 120
QY 121 LNRKNKPQIKIQKPKGKSELINRKASLADSGEYCKVSKLGNDSASANITIVESNEII 180
DB 121 LNRKNKPQIKIQKPKGKSELINRKASLADSGEYCKVSKLGNDSASANITIVESNEII 180
QY 181 TCMFASTEGAYVSSSPIRISVSTEGANTSSSTSTSTTCTSHLVKCAEKEKTCVNGGEC 240
DB 181 TCMFASTEGAYVSSSPIRISVSTEGANTSSSTSTSTTCTSHLVKCAEKEKTCVNGGEC 240
QY 241 FVVKDLSNPSRYLCKCPGFTGARTENVPKQVQKAEELYQKRVLTITGICITALLVY 300
DB 241 FVVKDLSNPSRYLCKCPGFTGARTENVPKQVQKAEELYQKRVLTITGICITALLVY 300
QY 301 GIMCVVAYCKTKQKRLKDLRLQSLRSERNNMNIANGPHHPNPPENVLVNOYVSKN 360
DB 301 GIMCVVAYCKTKQKRLKDLRLQSLRSERNNMNIANGPHHPNPPENVLVNOYVSKN 360
QY 361 VISSHIVEREAETSFSTSHYTSTAHHSTVTQTPSHWSNGHTESSILSESHSVIWMSSV 420
DB 361 VISSHIVEREAETSFSTSHYTSTAHHSTVTQTPSHWSNGHTESSILSESHSVIWMSSV 420
QY 421 ENSRHSSPTGGPRGLNGTGGPRECNSFLRHARETSDSYRDSPHSERYYSAMTTARMSP 480
DB 421 ENSRHSSPTGGPRGLNGTGGPRECNSFLRHARETSDSYRDSPHSERYYSAMTTARMSP 480
QY 481 VDFHTPSSPKSPPEMSPVSMTPSMAYSPFMEERPLLLVTPPLREKKKFDHHPQ 540
DB 481 VDFHTPSSPKSPPEMSPVSMTPSMAYSPFMEERPLLLVTPPLREKKKFDHHPQ 540
QY 541 QFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANSRRARTKPNNGHI 600
DB 541 QFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANSRRARTKPNNGHI 600
QY 601 ANRLEVDNSTSQSSNSESETEDEVRGDTPLGIONPLAASLEATPAFLADSTRTNAG 660
DB 601 ANRLEVDNSTSQSSNSESETEDEVRGDTPLGIONPLAASLEATPAFLADSTRTNAG 660
QY 661 RFTSQEEIQ 669
DB 661 RFTSQEEIQ 669
```

RESULT 3

US-09-792-025-1
; Sequence 1, Application US/09792025
; Patent No. US20020042087A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way
City: South San Francisco
State: California
Country: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/792,025
FILING DATE: 23-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/020,598
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Schwartz, Timothy R.
REGISTRATION NUMBER: 32171
REFERENCE/DOCKET NUMBER: P1145
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-7467
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
MOLECULE TYPE: amino acid
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-792-025-1

541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQETPEAQEPVKKLANRRARTRKPNNGHI 600
|||||
541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQETPEAQEPVKKLANRRARTRKPNNGHI 600
|||||
601 ANRLEVDNSTSSQSSNSETEDERVGDTPLGIONPLAASLEATPAFLADSTNPAG 660
|||||
601 ANRLEVDNSTSSQSSNSETEDERVGDTPLGIONPLAASLEATPAFLADSTNPAG 660
|||||
661 RFSTQEEIQ 669
|||||
661 RFSTQEEIQ 669
|||||
RESULT 4
US-09-849-868-1
; Sequence 1, Application US/09849868
; Patent No. US20020081299A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Gao, Wei-Qiang
; TITLE OF INVENTION: HAIR CELL DISORDERS
; FILE REFERENCE: GENENT.035VPC
; CURRENT APPLICATION NUMBER: US/09/849,868
; CURRENT FILING DATE: 2001-03-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-868-1

Query Match 100.0%; Score 3470; DB 10; Length 669;
Best Local Similarity 100.0%; Pred. No. 4e-173;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPGRSLSPSRDKLPNPTRALGPNSPAPRAVVRVSVSGENSERKEGKGKGGKK 60
DB 1 ARAPGRSLSPSRDKLPNPTRALGPNSPAPRAVVRVSVSGENSERKEGKGKGGKK 60
QY 61 ERGSKKPESAGSOSPALPPRLKEMKSOESAGSKLVLRCTSEYSSLRKFKWKNNE 120
DB 61 ERGSKKPESAGSOSPALPPRLKEMKSOESAGSKLVLRCTSEYSSLRKFKWKNNE 120
QY 121 LNRKNKPNQIKQKPKGKSELINKASLADSGEYMKVSKLGNDSASANITIVESNEII 180
DB 121 LNRKNKPNQIKQKPKGKSELINKASLADSGEYMKVSKLGNDSASANITIVESNEII 180
QY 181 TCMFASTGAYVSSSESPRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TCMFASTGAYVSSSESPRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FMVKDLSNPSRYLCKQPGFTGARTENVPKVNQOEAEELYQKRVLTITGICITALLV 300
DB 241 FMVKDLSNPSRYLCKQPGFTGARTENVPKVNQOEAEELYQKRVLTITGICITALLV 300
QY 301 GIMCVVAYCKTKKQKRLHDLRQSLRSERNMMNMIANGPHHPNPPNPVQVNVYVSKN 360
DB 301 GIMCVVAYCKTKKQKRLHDLRQSLRSERNMMNMIANGPHHPNPPNPVQVNVYVSKN 360
QY 361 VISSHHIYERAEATSFSTSHYTSTAHHSTTVTQTPSHWSNGHTSILSESHSVTVMSV 420
DB 361 VISSHHIYERAEATSFSTSHYTSTAHHSTTVTQTPSHWSNGHTSILSESHSVTVMSV 420
QY 421 ENSRHSSTGGPRGLNGTGGPRECNSFLRHARETPDSYRSPHSEYVVSAMTTPARMSP 480
DB 421 ENSRHSSTGGPRGLNGTGGPRECNSFLRHARETPDSYRSPHSEYVVSAMTTPARMSP 480
QY 481 VDFHTPSPKPPSPSPSPVSSMTYMPMSVSPMEERPLLLVTPPRLREKFKDHPHQ 540
DB 481 VDFHTPSPKPPSPSPSPVSSMTYMPMSVSPMEERPLLLVTPPRLREKFKDHPHQ 540

QY 1 ARAPGRSLSPSRDKLPNPTRALGPNSPAPRAVVRVSVSGENSERKEGKGKGGKK 60
DB 1 ARAPGRSLSPSRDKLPNPTRALGPNSPAPRAVVRVSVSGENSERKEGKGKGGKK 60
QY 61 ERGSKKPESAGSOSPALPPRLKEMKSOESAGSKLVLRCTSEYSSLRKFKWKNNE 120
DB 61 ERGSKKPESAGSOSPALPPRLKEMKSOESAGSKLVLRCTSEYSSLRKFKWKNNE 120
QY 121 LNRKNKPNQIKQKPKGKSELINKASLADSGEYMKVSKLGNDSASANITIVESNEII 180
DB 121 LNRKNKPNQIKQKPKGKSELINKASLADSGEYMKVSKLGNDSASANITIVESNEII 180
QY 181 TCMFASTGAYVSSSESPRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TCMFASTGAYVSSSESPRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FMVKDLSNPSRYLCKQPGFTGARTENVPKVNQOEAEELYQKRVLTITGICITALLV 300
DB 241 FMVKDLSNPSRYLCKQPGFTGARTENVPKVNQOEAEELYQKRVLTITGICITALLV 300
QY 301 GIMCVVAYCKTKKQKRLHDLRQSLRSERNMMNMIANGPHHPNPPNPVQVNVYVSKN 360
DB 301 GIMCVVAYCKTKKQKRLHDLRQSLRSERNMMNMIANGPHHPNPPNPVQVNVYVSKN 360
QY 361 VISSHHIYERAEATSFSTSHYTSTAHHSTTVTQTPSHWSNGHTSILSESHSVTVMSV 420
DB 361 VISSHHIYERAEATSFSTSHYTSTAHHSTTVTQTPSHWSNGHTSILSESHSVTVMSV 420
QY 421 ENSRHSSTGGPRGLNGTGGPRECNSFLRHARETPDSYRSPHSEYVVSAMTTPARMSP 480
DB 421 ENSRHSSTGGPRGLNGTGGPRECNSFLRHARETPDSYRSPHSEYVVSAMTTPARMSP 480
QY 481 VDFHTPSPKPPSPSPSPVSSMTYMPMSVSPMEERPLLLVTPPRLREKFKDHPHQ 540
DB 481 VDFHTPSPKPPSPSPSPVSSMTYMPMSVSPMEERPLLLVTPPRLREKFKDHPHQ 540
QY 541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQETPEAQEPVKKLANRRARTRKPNNGHI 600

Db 301 NVOLVNOYVSKNVISSHEIIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 360
QY 409 SESHVIVMSSVNSRHSSPTGGPRGLNGTGPRECNSFLRHARETPODSYRDSPHSERY 468
Db 361 SESHVIVMSSVNSRHSSPTGGPRGLNGTGPRECNSFLRHARETPODSYRDSPHS --Y 418
QY 469 VSAMTTPARMSPVDFTHTPPSPKSPPEMSPVSSMTVSMPSMAVSPFMEERPLLIVTTP 528
Db 419 VSAMTTPARMSPVDFTHTPPSPKSPPEMSPVSSMTVSMPSMAVSPFMEERPLLIVTTP 478
QY 529 RLREKFDHHPQOFSFHHPNPAHDSNLSPLASPLRIVEDEYETQYEPYPAQEPVKKLAN 588
Db 479 RLREKFDHHPQOFSFHHPNPAHDSNLSPLASPLRIVEDEYETQYEPYPAQEPVKKLAN 536
QY 589 RRAKRTKPNGHIANRLEVDNSTSSQSNSESETEDEVRGDTPLFGIONPLAASLEATPA 648
Db 537 RRAKRTKPNGHIANRLEVDNSTSSQSNSESETEDEVRGDTPLFGIONPLAASLEATPA 596
QY 649 FRLADSRTPNAGRFSTQEEIQ 669
Db 597 FRLADSRTPNAGRFSTQEEIQ 617

RESULT 9

US-09-795-668-16

; Sequence 16, Application US/09795668

; Patent No. US2002004557A1

; GENERAL INFORMATION:

; APPLICANT: Stefansson, Hreinn

; APPLICANT: Steinthorsdottir, Valgerdur

; APPLICANT: Gulcher, Jeffrey R.

; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE

; FILE REFERENCE: 2345.2004-001

; CURRENT APPLICATION NUMBER: US/09/795,668

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US 09/515,716

; NUMBER OF SEQ ID NOS: 1531

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 632

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-795-668-16

Query Match 90.9%; Score 3153; DB 10; Length 632;

Best Local Similarity 98.7%; Pred. No. 1.2e-156;

Matches 613; Conservative 0; Mismatches 2; Indels 6; Gaps 3;

49 KEGRGKGGKKGKGGKGPESAGSOSPALPRLKEMKSOESAAGSKVLRCETSEYS 108
Db 3 KEGRGKGGKKGKGGKGPESAGSOSPALPRLKEMKSOESAAGSKVLRCETSEYS 62

109 SLRFKFWKNGNENLNKKNQIKQKPKSELINKASLADSGEYMKVSKLGNDSAS 168
Db 63 SLRFKFWKNGNENLNKKNQIKQKPKSELINKASLADSGEYMKVSKLGNDSAS 122

169 ANTIIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAE 228
Db 123 ANTIIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAE 182

229 KERTFCVNGEGCFMWKDLNPSRYLCKOPGTGARTENVPKMQNQKAEELYOKRYL 288
Db 183 KERTFCVNGEGCFMWKDLNPSRYLCKOPGTGARTENVPKMQNQKAEELYOKRYL 242

289 TITGICALLVGMICVWAYCKTKQKKLHDLRLQSLRSERNMNIANGPHHPNPPPE 348
Db 243 TITGICALLVGMICVWAYCKTKQKKLHDLRLQSLRS --NNTNANIANGPHHPNPPPE 300

349 NVOLVNOYVSKNVISSHEIIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 408
QY 301 NVOLVNOYVSKNVISSHEIIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 360

QY 409 SESHVIVMSSVNSRHSSPTGGPRGLNGTGPRECNSFLRHARETPODSYRDSPHSERY 468
Db 361 SESHVIVMSSVNSRHSSPTGGPRGLNGTGPRECNSFLRHARETPODSYRDSPHS --Y 418
QY 469 VSAMTTPARMSPVDFTHTPPSPKSPPEMSPVSSMTVSMPSMAVSPFMEERPLLIVTTP 528
Db 419 VSAMTTPARMSPVDFTHTPPSPKSPPEMSPVSSMTVSMPSMAVSPFMEERPLLIVTTP 478
QY 529 RLREKFDHHPQOFSFHHPNPAHDSNLSPLASPLRIVEDEYETQYEPYPAQEPVKKLAN 588
Db 479 RLREKFDHHPQOFSFHHPNPAHDSNLSPLASPLRIVEDEYETQYEPYPAQEPVKKLAN 536
QY 589 RRAKRTKPNGHIANRLEVDNSTSSQSNSESETEDEVRGDTPLFGIONPLAASLEATPA 648
Db 537 RRAKRTKPNGHIANRLEVDNSTSSQSNSESETEDEVRGDTPLFGIONPLAASLEATPA 596
QY 649 FRLADSRTPNAGRFSTQEEIQ 669
Db 597 FRLADSRTPNAGRFSTQEEIQ 617

RESULT 10

US-09-795-686-16

; Sequence 16, Application US/09795686

; Patent No. US20020094954A1

; GENERAL INFORMATION:

; APPLICANT: Stefansson, Hreinn

; APPLICANT: Steinthorsdottir, Valgerdur

; APPLICANT: Gulcher, Jeffrey R.

; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE

; FILE REFERENCE: 2345.2005-001

; CURRENT APPLICATION NUMBER: US/09/795,686

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US 09/515,715

; NUMBER OF SEQ ID NOS: 1531

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 632

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-795-686-16

Query Match 90.9%; Score 3153; DB 10; Length 632;

Best Local Similarity 98.7%; Pred. No. 1.2e-156;

Matches 613; Conservative 0; Mismatches 2; Indels 6; Gaps 3;

49 KEGRGKGGKKGKGGKGPESAGSOSPALPRLKEMKSOESAAGSKVLRCETSEYS 108
Db 3 KEGRGKGGKKGKGGKGPESAGSOSPALPRLKEMKSOESAAGSKVLRCETSEYS 62

109 SLRFKFWKNGNENLNKKNQIKQKPKSELINKASLADSGEYMKVSKLGNDSAS 168
Db 63 SLRFKFWKNGNENLNKKNQIKQKPKSELINKASLADSGEYMKVSKLGNDSAS 122

169 ANTIIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAE 228
Db 123 ANTIIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAE 182

229 KERTFCVNGEGCFMWKDLNPSRYLCKOPGTGARTENVPKMQNQKAEELYOKRYL 288
Db 183 KERTFCVNGEGCFMWKDLNPSRYLCKOPGTGARTENVPKMQNQKAEELYOKRYL 242

289 TITGICALLVGMICVWAYCKTKQKKLHDLRLQSLRSERNMNIANGPHHPNPPPE 348
Db 243 TITGICALLVGMICVWAYCKTKQKKLHDLRLQSLRS --NNTNANIANGPHHPNPPPE 300

349 NVOLVNOYVSKNVISSHEIIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 408
QY 301 NVOLVNOYVSKNVISSHEIIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 360

409 SESHVIVMSSVNSRHSSPTGGPRGLNGTGPRECNSFLRHARETPODSYRDSPHSERY 468
QY

Db 361 SESHSIVMSSVENSRRHSSPTGGPRGRNGTGGPRECNFSLRHARETDPDSTRDPSHS--Y 418
Qy 469 VSAMTTPARMSPVDFHTPSSPKSPSEMSPVSSMTVSMPSMAYSPFMEERPLLLVTPP 528
Db 419 VSAMTTPARMSPVDFHTPSSPKSPSEMSPVSSMTVSMPSMAYSPFMEERPLLLVTPP 478
Qy 529 RLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEEYETTOEYEPAPQEPVKKLANS 588
Db 479 RLREKKFDHHPQOFSFHHNPAHDSNLSPL--DLRIVEDEEYETTOEYEPAPQEPVKKLANS 536
Qy 589 RRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDEVRGDTPLGTONPLAASLEATPA 648
Db 537 RRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDEVRGDTPLGTONPLAASLEATPA 596
Qy 649 FRLADSRTPNAGRFSTQBEIQ 669
Db 597 FRLADSRTPNAGRFSTQBEIQ 617

RESULT 11

US-10-096-241-10
Sequence 10, Application US/10096241
Patent No. US20020127594A1
GENERAL INFORMATION:
APPLICANT: Gearling, David P.
TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
AND USES THEREFOR
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/096,241
FILING DATE: 12-Mar-2002
CLASSIFICATION DATA:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/699,591
FILING DATE: 19-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07334/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 645 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-096-241-10

Query Match 90.38; Score 3133.5; DB 12; Length 645;
Best Local Similarity 96.78; Pred. No. 1.2e-155;
Matches 609; Conservative 1; Mismatches 15; Indels 5; Gaps 1;
Qy 45 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAAGSKLVLRCTS 104
Db 1 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAAGSKLVLRCTS 60

Qy 105 SEYSSLRFRKFKNGNLRNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 164
Db 61 SEYSSLRFRKFKNGNLRNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120
Qy 165 DSASANTIVESNEIITGMPASTEGAYVSSSPTRISVSTEGANTSSSTSTTGTSHLV 224
Db 121 DSASANTIVESNEIITGMPASTEGAYVSSSPTRISVSTEGANTSSSTSTTGTSHLV 180
Qy 225 KCAKEKTFVNGGCEFMVDLSNPRLCKQPGFTGACATENVPKVNQ-----EKA 279
Db 181 KCAKEKTFVNGGCEFMVDLSNPRLCKQPGFTGACATENVPKVNQ-----EKA 240
Qy 280 EELYQKRVLTITGICIAALLVVGIMCVVAYCKTKKQKRLHDLRQLSRERNMMNIANG 339
Db 241 EELYQKRVLTITGICIAALLVVGIMCVVAYCKTKKQKRLHDLRQLSRERNMMNIANG 300
Qy 340 PHHPNPPENVQLVNYVSKNVISSEHIVREAEATSESTSHYSTAHHSTVTTQTPSHSW 399
Db 301 PHHPNPPENVQLVNYVSKNVISSEHIVREAEATSESTSHYSTAHHSTVTTQTPSHSW 360
Qy 400 SNHTESILSESHSVIVMSSVENSRRHSPTGGPRGRNGTGGPRECNFSLRHARETDPDY 459
Db 361 SNHTESILSESHSVIVMSSVENSRRHSPTGGPRGRNGTGGPRECNFSLRHARETDPDY 420
Qy 460 RDSPHSERVYSAMTTPARMSPVDFHTPSSPKSPSEMSPVSSMTVSMPSMAYSPFMEEE 519
Db 421 RDSPHSERVYSAMTTPARMSPVDFHTPSSPKSPSEMSPVSSMTVSMPSMAYSPFMEEE 480
Qy 520 RPLLLVTPRLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEEYETTOEYEPAPQ 579
Db 481 RPLLLVTPRLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEEYETTOEYEPAPQ 540
Qy 580 EPVKKLANSRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDEVRGDTPLGIONPL 639
Db 541 EPVKKLANSRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDEVRGDTPLGIONPL 600
Qy 640 AASLEATPAPRLADSRTPNAGRFSTQBEIQ 669
Db 601 AASLEATPAPRLADSRTPNAGRFSTQBEIQ 630

RESULT 12

US-09-773-517-5
Sequence 5, Application US/09773517
Patent No. US20010023241A1
GENERAL INFORMATION:
APPLICANT: Sliwowski, Mark
APPLICANT: Kern, Jeffrey
TITLE OF INVENTION: Use of Heregulin as a Growth Factor
FILE REFERENCE: P1145R1
CURRENT APPLICATION NUMBER: US/09/773,517
CURRENT FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 09/243,198
PRIOR FILING DATE: 1999-02-02
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 5
LENGTH: 637
TYPE: PRT
ORGANISM: Homo sapiens
US-09-773-517-5

Query Match 90.28; Score 3131.5; DB 10; Length 637;
Best Local Similarity 97.38; Pred. No. 1.5e-155;
Matches 608; Conservative 3; Mismatches 11; Indels 3; Gaps 1;

Qy 45 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAAGSKLVLRCTS 104
Db 1 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAAGSKLVLRCTS 60
Qy 105 SEYSSLRFRKFKNGNLRNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 164
Db 61 SEYSSLRFRKFKNGNLRNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120

[illegible]

RESULT 13

US-09-792-025-5
Sequence 5, Application US/09792025
Patent No. US20020042087A1
GENERAL INFORMATION:
APPLICANT: Sliwowski, Mark
Kern, Jeffrey
TITLE OF INVENTION: Use of Heregulin as a Growth Factor
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/792,025
FILING DATE: 23-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/020,598
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Schwartz, Timothy R.
REGISTRATION NUMBER: 32171
REFERENCE/DOCKET NUMBER: P1145
TELEPHONE: 650/225-7467
TELEFAX: 650/952-9881

```

RESULT 14
US-09-849-868-5
; Sequence 5, Application US/09849868
; Patent No. US20020081299A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Gao, Wei-Qiang
; TITLE OF INVENTION: HAIR CELL DISORDERS
; FILE REFERENCE: GENENT.035VPC
; CURRENT APPLICATION NUMBER: US/09/849,868
; CURRENT FILING DATE: 2001-05-04
; . NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 637

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-868-5

```

Query Match 90.2%; Score 3131.5; DB 10; Length 637;
Best Local Similarity 97.3%; Pred. No. 1.5e-155;
Matches 608; Conservative 3; Mismatches 11; Indels 3; Gaps 1;

Qy	45	MSRKEBGRGKGKGGKGGSGKKPESAAAGSOSPAPPLKEMKQESAAAGSKVLJRCSTS	104
Db	1	MSRKEBGRGKGKGGKGGSGKKPESAAAGSOSPAPPLKEMKQESAAAGSKVLJRCSTS	60
Qy	105	SEYSSSLRFKFKNGNELNRKNKPNIOIKQKPGKSELINKASLADSGEYCKVLSKIGN	164
Db	61	SEYSSSLRFKFKNGNELNRKNKPNIOIKQKPGKSELINKASLADSGEYCKVLSKIGN	120
Qy	165	DSASANTIVESNEIITGMPASTEGAYVSSPISIRISVSTEGANTSSSTSTSTTGTSHLV	224
Db	121	DSASANTIVESNEIITGMPASTEGAYVSSPISIRISVSTEGANTSSSTSTSTTGTSHLV	180
Qy	225	KCAEKEKTFCVNGGCEFMVLDLSPNSRYLCKQCFPGCARCTENVPMKVQOEKAEELYQ	284
Db	181	KCAEKEKTFCVNGGCEFMVLDLSPNSRYLCKQCFPGCARCTENVPMKVQOEKAEELYQ	237
Qy	285	KRVLTITGICIALLVGIMCVACKYKQKORKLHDLRQSLRSERNNMNIANGPHPN	344
Db	238	KRVLTITGICIALLVGIMCVACKYKQKORKLHDLRQSLRSERNNMNIANGPHPN	297
Qy	345	PPPENVLVNOQYVSKNVISSEHIVERBAEYFSFSTSHYTSTAHSSTVTQTSPSHWSNGHT	404
Db	298	PPPENVLVNOQYVSKNVISSEHIVERBAEYFSFSTSHYTSTAHSSTVTQTSPSHWSNGHT	357
Qy	405	ESILSESHSVTVMSVSVNSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYRDSPH	464
Db	358	ESILSESHSVTVMSVSVNSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYRDSPH	417
Qy	465	SERYVSAMTTPARMSPVDFHTPSPKPPSPSEMSPPVSSMTVSPMSVSPPEERPLLL	524
Db	418	SERYVSAMTTPARMSPVDFHTPSPKPPSPSEMSPPVSSMTVSPMSVSPPEERPLLL	477
Qy	525	VTPLRREKFFDHHPOQFSSFHHPNADHNSLPASPLRIVEDEYETQOYEPQEPVKK	584
Db	478	VTPLRREKFFDHHPOQFSSFHHPNADHNSLPASPLRIVEDEYETQOYEPQEPVKK	537
Qy	585	LANSRRAKRTKPNGHIANRLVDSDNTSSQSSNSESETEDERVGEDTTPGLIGNPLAASLE	644
Db	538	LANSRRAKRTKPNGHIANRLVDSDNTSSQSSNSESETEDERVGEDTTPGLIGNPLAASLE	597
Qy	645	ATPAFRLADSRNTNPAGRFSTQEETQ	669
Db	598	ATPAFRLADSRNTNPAGRFSTQEETQ	622

RESULT 15

```

RES001.12
US-09-946-807-14
; Sequence 14, Application US/09946807
; Patent No. US20020165144A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthoroddottir, Valgerdur
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/946.807
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 629
;

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-946-807-14

```

Query Match 87.8%; Score 3047.5; DB 9; Length 629;
Best Local Similarity 96.3%; Pred. No. 3.6e-151;
Matches 598; Conservative 2; Mismatches 12; Indels 9; Gaps 4

QY	49	KEGRGKGGKKKRGSGKKP	PSAAGSQSPALPPRLKEMKQESAAAGSKLVR	CETSEYS	108
Db	3	KEGRGKGGKKKRGSGKKP	SAAGSQSPALPPRLKEMKQESAAAGSKLVR	CETSEYS	62
QY	109	SLRFKWFKNGLNKNPNQIK	QKPKGKSELINKASLADSGEYMKV	ISKLGNDAS	168
Db	63	SLRFKWFKNGLNKNPNQIK	QKPKGKSELINKASLADSGEYMKV	ISKLGNDAS	122
QY	169	ANITIVESNEITGMPASTE	GAYVSSSEPIRISYSTEGANTSSSTST	TGTHSLVKCAE	228
Db	123	ANITIVESNEITGMPASTE	GAYVSSSEPIRISYSTEGANTSSSTST	TGTHSLVKCAE	182
QY	229	KEKTFCVNGGECFMVKD	LSNPSRYLCKQCPGTGARTENVPMKVQ	QKAEELYOKRVL	288
Db	183	KEKTFCVNGGECFMVKD	LSNPSRYLCKQCPNFTGDRQNTYV	--MASFYKAEELYOKRVL	239
QY	289	TITGICIALLVGIMCVWAY	CKTKQRKKLHDLRQSLRBERNNMNIANG	PHHPPPPE	348
Db	240	TITGICIALLVGIMCVWAY	CKTKQRKKLHDLRQSLRS	--NNTMNIANGPHHPPPPE	297
QY	349	NVOLNQYYSKVNVISSE	HIVEREAFTSFSTSHYTSTAHHSTTVTQTP	SHSWSNGHTESIL	408
Db	298	NVOLNQYYSKVNVISSE	HIVEREAFTSFSTSHYTSTAHHSTTVTQTP	SHSWSNGHTESIL	357
QY	409	SESHSVIVNSSVENSRRH	SPGGRGLNGTGGPRECNSFLRHARET	PDOSYRDSPHERY	468
Db	358	SESHSVIVNSSVENSRRH	SPGGRGLNGTGGPRECNSFLRHARET	PDOSYRDSPHS	--Y 415
QY	469	VSAMTTPARMSVDFTHT	SPSPKSPPEMSPPVSSMTVSMPSMAV	SFPMEEERPLLVTTP	528
Db	416	VSAMTTPARMSVDFTHT	SPSPKSPPEMSPPVSSMTVSMPSMAV	SFPMEEERPLLVTTP	475
QY	529	RLREKFFDHHPOQFSF	FHNPAHDSNSLPASPLRIVEDEYEY	TQYEPAPQEPVKLIANS	588
Db	476	RLREKFFDHHPOQFSF	FHNPAHDSNSLP--DLRIVEDEYEY	TQYEPAPQEPVKLIANS	533
QY	589	RAKRTKPNGHIANLEVD	SNTSQSSNSSETEDEVRGDTPLG	TQNPPLAASLEATPA	648
Db	534	RAKRTKPNGHIANLEVD	SNTSQSSNSSETEDEVRGDTPLG	TQNPPLAASLEATPA	593
QY	649	FRLADSRTPNAGRFST	QBEIQ 669		
Db	594	FRLADSRTPNAGRFST	QBEIO 614		

Search completed: July 9, 2003, 10:49:15
Job time : 30 secs

Result No.	Query			DB ID	Description
	Score	Match	Length		
1	3243	93.5	640	2	A43273
2	3135.5	90.4	637	2	C43273
3	3133.5	90.3	645	2	B43273
4	3021.5	87.1	639	2	I61719
5	2922	84.2	636	2	I61718
6	2917	84.1	662	2	I61722
7	2314	66.7	602	2	A45769
8	2204	63.5	462	2	I38404
9	1804	52.0	350	2	I38403
10	1109	32.0	241	2	D43273
11	1024	29.5	241	2	S32359
12	1013.5	29.2	850	2	JC5700
13	1012.5	29.2	860	2	JC5702
14	1001.5	28.9	868	2	JC5701
15	985	28.4	230	2	A56210
16	750.5	21.6	422	2	S32357
17	710	20.5	175	2	I38408
18	598.5	17.2	125	2	S62676
19	581	16.7	125	2	I38405
20	301	8.7	713	2	T44447
21	273	7.9	296	2	A56943
22	180.5	5.2	5327	2	T13564
23	179.5	5.2	2254	2	D86215
24	176	5.1	734	2	B42680
25	172	5.0	3507	2	T34513
26	170	4.9	990	2	I51618
27	166.5	4.8	1459	2	T32271
28	166.5	4.8	1510	2	T33100
29	165	4.8	1091	1	I3CHN1

QY 165 DSANANITIVESNEIITGMPASTGAGYVSSESPIRISYSTEGANTSSSTSTSTTGTSHLV 224
|||||
Db 121 DSANANITIVESNEIITGMPASTGAGYVSSESPIRISYSTEGANTSSSTSTSTTGTSHLV 180
|||||
QY 225 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQPGFTGARCFTENPMKVONOEKAEELYQ 284
|||||
Db 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQPGFTGARCFTENPMKVONOEKAEELYQ 240
|||||
QY 285 KRVLTITGICIAALLVVGIMCVVAYCKTKKORKKLDRLRQSLRSERNMMNIANGPHHPN 344
|||||
Db 241 KRVLTITGICIAALLVVGIMCVVAYCKTKKORKKLDRLRQSLRSERNMMNIANGPHHPN 300
|||||
QY 345 PPNENVOLVQVYKNSVISEHIVERAEATSFSTSHYTSTAHHSTVTTQTPSHSWSNGHT 404
|||||
Db 301 PPNENVOLVQVYKNSVISEHIVERAEATSFSTSHYTSTAHHSTVTTQTPSHSWSNGHT 360
|||||
QY 405 ESILSESHSVIYVSSVENSRRHSSPTGGPRGRNGTGGPRECNFSFLRHARETSDSYRDSPH 464
|||||
Db 361 ESILSESHSVIYVSSVENSRRHSSPTGGPRGRNGTGGPRECNFSFLRHARETSDSYRDSPH 420
|||||
QY 465 SERYVSAMTTPARMSPVDFHTPSSPKSPSPSEMSPVSSMTVSMPSMAYSPFMEERPLLL 524
|||||
Db 421 SERYVSAMTTPARMSPVDFHTPSSPKSPSPSEMSPVSSMTVSMPSMAYSPFMEERPLLL 480
|||||
QY 525 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYETTOEYEPAPQPVKK 584
|||||
Db 481 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYETTOEYEPAPQPVKK 540
|||||
QY 585 LANSRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGEDTDFLGIQNPAAASLE 644
|||||
Db 541 LANSRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGEDTDFLGIQNPAAASLE 600
|||||
QY 645 ATPAPRLADSRTPNAGRFSTQEEIQ 669
|||||
Db 601 ATPAPRLADSRTPNAGRFSTQEEIQ 625
|||||
RESULT 2
C43273
heregulin precursor, splice form beta-2 - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Aug-2002
C:Accession: C43273; I38407
R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu
Science 256, 1205-1210, 1992
A:Title: Identification of heregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253; PMID:1350381
A:Accession: C43273
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-637 <HOL>
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863; PMID:7509448
A:Accession: I38407
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 119-406 <RES>
A:Cross-references: EMBL:U02329; NID:g408408; PIDN:AAA19954.1; PID:g408409
C:Genetics:
A:Gene: GDB:HGL
A:Cross-references: GDB:I32656; OMIM:142445
A:Map position: 8p22-8p11
C:Superfamily: human heregulin; EGF homology
C:Keywords: alternative splicing
F:182-221/Domain: EGF homology <EGF>
Query Match 90.4%; Score 3135.5; DB 2; Length 637;
Best Local Similarity 97.4%; Pred. No. 2.2e-167;
Matches 609; Conservative 2; Mismatches 11; Indels 3; Gaps 1;
QY 45 MSEREGRGKGGKKGKPSAAGSQSPALPPLKEMKSOESAGSKLVLRCTS 104

Db 1 MSEREGRGKGGKKGKPSAAGSQSPALPPLKEMKSOESAGSKLVLRCTS 60
|||||
QY 105 SEYSLRKPWPKNGNELNRKNKPONIKQKPKGSELRLINKASLADSGYCMCKVLSKLG 164
|||||
Db 61 SEYSLRKPWPKNGNELNRKNKPONIKQKPKGSELRLINKASLADSGYCMCKVLSKLG 120
|||||
QY 165 DSANANITIVESNEIITGMPASTGAGYVSSESPIRISYSTEGANTSSSTSTSTTGTSHLV 224
|||||
Db 121 DSANANITIVESNEIITGMPASTGAGYVSSESPIRISYSTEGANTSSSTSTSTTGTSHLV 180
|||||
QY 225 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQPGFTGARCFTENPMKVONOEKAEELYQ 284
|||||
Db 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQPGFTGARCFTENPMKVONOEKAEELYQ 237
|||||
QY 285 KRVLTITGICIAALLVVGIMCVVAYCKTKKORKKLDRLRQSLRSERNMMNIANGPHHPN 344
|||||
Db 238 KRVLTITGICIAALLVVGIMCVVAYCKTKKORKKLDRLRQSLRSERNMMNIANGPHHPN 297
|||||
QY 345 PPNENVOLVQVYKNSVISEHIVERAEATSFSTSHYTSTAHHSTVTTQTPSHSWSNGHT 404
|||||
Db 298 PPNENVOLVQVYKNSVISEHIVERAEATSFSTSHYTSTAHHSTVTTQTPSHSWSNGHT 357
|||||
QY 405 ESILSESHSVIYVSSVENSRRHSSPTGGPRGRNGTGGPRECNFSFLRHARETSDSYRDSPH 464
|||||
Db 358 ESILSESHSVIYVSSVENSRRHSSPTGGPRGRNGTGGPRECNFSFLRHARETSDSYRDSPH 417
|||||
QY 465 SERYVSAMTTPARMSPVDFHTPSSPKSPSPSEMSPVSSMTVSMPSMAYSPFMEERPLLL 524
|||||
Db 418 SERYVSAMTTPARMSPVDFHTPSSPKSPSPSEMSPVSSMTVSMPSMAYSPFMEERPLLL 477
|||||
QY 525 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYETTOEYEPAPQPVKK 584
|||||
Db 478 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYETTOEYEPAPQPVKK 537
|||||
QY 585 LANSRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGEDTDFLGIQNPAAASLE 644
|||||
Db 538 LANSRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGEDTDFLGIQNPAAASLE 597
|||||
QY 645 ATPAPRLADSRTPNAGRFSTQEEIQ 669
|||||
Db 598 ATPAPRLADSRTPNAGRFSTQEEIQ 622
|||||
RESULT 3
B43273
heregulin, splice form beta 1 - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Aug-2002
C:Accession: B43273; I38406
R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yan
Science 256, 1205-1210, 1992
A:Title: Identification of heregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253; PMID:1350381
A:Accession: B43273
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-645 <HOL>
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation
A:Reference number: A56210; MUID:94158863; PMID:7509448
A:Accession: I38406
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 'A', 95-418, 'F', 420-645 <RES>
A:Cross-references: EMBL:U02328; NID:g408406; PIDN:AAA19953.1; PID:g408407
C:Genetics:
A:Gene: GDB:HGL
A:Cross-references: GDB:I32656; OMIM:142445
A:Map position: 8p22-8p11
C:Superfamily: human heregulin; EGF homology
C:Keywords: alternative splicing

Mol. Cell. Biol. 14, 1909-1919, 1994

A;Title: Structural and functional aspects of the multiplicity of Neu differentiation

A;Reference number: A56210; MUID:94158863; PMID:7509448

A;Accession: I38404

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-462 <RES>

A;Cross-references: EMBL:U02326; NID:g408402; PIDN:AAA19951.1; PID:g408403

C;Superfamily: human heregulin; EGF homology

Query Match 63.5%; Score 2204; DB 2; Length 462;

Best Local Similarity 99.8%; Pred. No. 1.2e-115;

Matches 423; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 45 MSERKEGRGKGKGGKRGSGKKPESAGSQSPALPRLKEMKQSQAAGSKLVLRCEYS 104

DB 1 MSERKEGRGKGKGGKRGSGKKPESAGSQSPALPRLKEMKQSQAAGSKLVLRCEYS 60

QY 105 SEYSSSLRFKWFKNGLNRKNKPNQIKKKPKSELINKASLADSGYMKVKVSKLGN 164

DB 61 SEYSSSLRFKWFKNGLNRKNKPNQIKKKPKSELINKASLADSGYMKVKVSKLGN 120

QY 165 DSASANTIVESNEIITGMPASTGAGVAYSESPIRISVTEGANTSSSTSTGTGSHLV 224

DB 121 DSASANTIVESNEIITGMPASTGAGVAYSESPIRISVTEGANTSSSTSTGTGSHLV 180

QY 225 KCAEKERTFCVNGGECFVKWDLNPNRYLCKCPGFTGARTENPMKVQNOEKAEELQ 284

DB 181 KCAEKERTFCVNGGECFVKWDLNPNRYLCKCPGFTGARTENPMKVQNOEKAEELQ 240

QY 285 KRVLTITGICIAALLVVGIMCVWAYCKTKQKKLHDLRQSLRSERNMMNTANGPHHPN 344

DB 241 KRVLTITGICIAALLVVGIMCVWAYCKTKQKKLHDLRQSLRSERNMMNTANGPHHPN 300

QY 345 PPNENVOLVQYYSKNVVISSEHIVERAETSSTSHYTSTAHHSTTVTTPSHSWSNGHT 404

DB 301 PPNENVOLVQYYSKNVVISSEHIVERAETSSTSHYTSTAHHSTTVTTPSHSWSNGHT 360

QY 405 ESILSESHSVIVMSSVENSRRHSSPTGPRGLNGTGGPRECNSFLRHARETDPDSYRDSPH 464

DB 361 ESILSESHSVIVMSSVENSRRHSSPTGPRGLNGTGGPRECNSFLRHARETDPDSYRDSPH 420

QY 465 SERY 468

DB 421 SERH 424

RESULT 9

I38403

C;Species: Homo sapiens (man)

C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 02-Aug-2002

C;Accession: I38403

R;Wen, D.; Suggs, S.V.; Karungaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.

Mol. Cell. Biol. 14, 1909-1919, 1994

A;Title: Structural and functional aspects of the multiplicity of Neu differentiation

A;Reference number: A56210; MUID:94158863; PMID:7509448

A;Accession: I38403

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-350 <RES>

A;Cross-references: EMBL:U02325; NID:g408400; PIDN:AAA19950.1; PID:g408401

C;Superfamily: human heregulin; EGF homology

Query Match 52.0%; Score 1804; DB 2; Length 350;

Best Local Similarity 97.4%; Pred. No. 1.6e-93;

Matches 341; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 238 GECFWKDLNPNRYLCKCPGFTGARTENPMKVQNOEK-----AEELYOKRVLT 289

DB 1 GECFWKDLNPNRYLCKCPGFTGARTENPMKVQNOEKHLGIFAEELYOKRVLT 60

QY 290 ITGICIAALLVVGIMCVWAYCKTKQKKLHDLRQSLRSERNMMNTANGPHHPNPPPN 349

[illegible]

RESULT 13
JC5702

Erbb kinase activator alpha2a, brain and thymus - rat
 Species: Rattus norvegicus (Norway rat)
 Date: 25-Nov-1997 #sequence_revision 25-Nov-1997 #text_change 02-Aug-2002
 Accession: JC5702; PC4417
 R;Hqashiyama, S.; Horkawa, M.; Yamada, K.; Ichino, N.; Nakano, N.; Nakagawa, T.; Miyag
 J. Biochem. 122, 673-680, 1997
 A;Title: A novel brain-derived member of the epidermal growth factor family that interact
 A;Reference number: JC5700; MUID:98006324; PMID:9348101
 A;Accession: JC5702
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-860 <HIG>
 A;Cross-references: DDBJ:D89996; NID:g2605631; PIDN:BAA23345.1; PID:g2605632
 A;Experimental source: PC-12 cell
 A;Accession: PC4417
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-212-213, 223-860 <HIG>
 A;Cross-references: DDBJ:AB001576; NID:g2605478; PIDN:BAA23348.1; PID:g2605479
 A;Experimental source: PC-12 cell
 C;Comment: This protein is a member of the epidermal growth factor family. It is function
 ing the differentiation of MDA-MB-453 cells.
 C;Superfamily: human Erbb kinase activator alpha, brain and thymus; EGF homology
 C;Keywords: glycoprotein
 F:274-327/Domain: Ig-like #status predicted <IGL>
 F:361-397/Domain: EGF homology <EGF>

F:422-444/Domain: hydrophobic #status predicted <HD>
F:163,294,467/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	29.2%;	Score 1012.5;	DB 2;	Length 860;
Best Local Similarity	37.4%;	Pred. No. 6.2e-49;		
Matches 249;	Conservative 102;	Mismatches 183;	Indels 131;	Gaps 21;

QY 55 GKKKRGGSGKKPESAGSQSPALPRLEKMKSQESAAGSKLVLCRTSSSYSSLSRFKW 114
Db GKNLKKEVGKILTCAT-----RPLKKMKSGTGEVGEKQSLACEAAGNPQPSYRW 286
QY 115 FKNGNELNRKNKQNTIKQKPK--SELINKASLADSGEYMKVYISKLNDASANT 172
Db FKDGKELNRG---RDIRIKNGRKNSRLQFNKRVKVEDAGEYCEAEITGLKDTVRGLH 343
QY 173 IVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTSTGTGTHLVKCAKEKT 232
Db 344 V-----NSVETLSSWSG--HARKCNETAKS 367
QY 233 FCVNGECFMVKDLSNPRLCYLCKQCGFTGARTENVMKV---QOKEAEELYQKRVLT 289
Db YCVNGVCYVYIEGINKLS---CKCPNGFFQRCLEKPLRLYMPDPKQAEELYQKRVLT 424
QY 290 ITGICALLVVGIMCVVAYCKTKQKKLHDLRLQSLRSERNNMNMIANGPHHPPPPP 349
Db ITGICVALLVVGIVCVVAYCKTKKQRQMHHLRQNMCPAHQN--RSLANGPSHRLDPEE 483
QY 350 VQLVNOVSKNVISSSEHIVERAEISFSTSHYTSTAHHSTTVTOT-----PSHSWSNGHT 404
Db IQWAD-YISKNVPATDVIYREAEITFSGSHSCSPSHHCSTATPTSSHRHESHHTWSLERS 542
QY 405 ESILSESHSVYVMSVSNSSHSSPT--GGPRGLRNGTGGPRECNFSFURHARETP-----DS 458
Db ESLTSDSQGIMLSSVGTSCNCPACVEARARAAAYSQEE-----RRRAAMPYHDSIDS 598
QY 459 YRSPHSIRYSVAMTPARMSPVDFHTPSSPKPPPEMGPVSSMTVSNPMSA--VSPFME 517
Db LRSPHSIRYSVAMTPARMSPVDFHTPSSPKPPPEMGPVSSMTVSNPMSA--VSPFME 517
QY 518 EERPLLVLTPRLREKKFDHP-----QOESSFPHNPAHDS-----553
Db EQPPLL-----RHPAPPGPGPGGADMQRSYDSYYPAGGPRGACALGG 705
QY 554 --NSLPASPLRIVEDEYEYTOEYPAQEVKKL-----ANSRRAKRTKPNGHIANR 603
Db SLGSLPASLPRIPEDEYEYTOEYPAQEVKKL-----ANSRRAKRTKPNGHIANR 603
QY 604 -----LEVDSNTSQSSNSESETERVG-----EDTFGLIQNPL--ASLETPAFRL 651
Db ARAARDLSLSSGSGCSASASDDADDADGALAAESTPTFLGLRAAHDLRSDSPPLCPA 825
QY 652 ADSRT 656
Db 826 ADSRT 830

RESULT 14
JC5701
Erbb kinase activator alphas, brain and thymus - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Nov-1997 #sequence_revision 25-Nov-1997 #text_change 02-Aug-2002
C:Accession: JC5701; PC4411
R:Higashiyama, S.; Horikawa, M.; Yamada, K.; Ichino, N.; Nakano, N.; Nakagawa, T.
J. Biochem. 122, 675-680, 1997
A:Title: A novel brain-derived member of the epidermal growth factor family that
A:Reference number: JC5700; PMID:98006324; PMID:9348101
A:Accession: JC5701
A:Molecule type: mRNA
A:Residues: 1-868 <HIG>
A:Cross-references: DDBJ:D89995; NID:g2605629; PIDN:BAA23344.1; PTD:g2605630
A:Accession: PC4411
A:Molecule type: protein
A:Residues: 128-162 <HIG>

A:Experimental source: PC-12 cell
C:Comment: This protein is a member of the epidermal growth factor family. It is functioning the differentiation of MDA-MB-453 cells.
C:Superfamily: human ErbB kinase activator alpha, brain and thymus; EGF homology F;361-397/Domain: EGF homology <EGF>

```
Query Match      28.9%; Score 1001.5; DB 2; Length 868;
Best Local Similarity 36.8%; Pred. No. 2.6e-48;
Matches 248; Conservative 102; Mismatches 184; Indels 139; Gaps 21;

QY 55 GKKKKRGSGKKPESAGSQSPALPPRLKEMKQESAKLVLCETSSSEYSLRFKW 114
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 234 GNKIKVEKILCTDCAT-----RPKLKMKKSGTGEVGEKQSLACEAANPQPSYRW 286
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 115 FRNGNELNRKNPKQNIQKPKG--SELIRKASLADSGEYMKVSKLGNDSASANIT 172
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 287 FKDGKELNRS--RDIRIKYNGRKNSRLQFNKVKVEDAGEYVCEAEINLGRDVGRLH 343
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 173 IVESNELITGMPASTGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKT 232
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
344 V-----NSVSTTLSSWSG--HARRKNETAKS 367

QY 233 FCVNGGECFMVKDLSNPSRYLCKCPGFTGARCTENVPMKVQ-----NOEKAEE 281
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 368 YCVNGGVCIYIEGINOLS---CKCPNGFFGQRCLEKPLRLYMPDPKQKHLGFELKEAE 424
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 282 LYQKRVLTITGICIAALLVVGIMCVWAYCKTKQKKLHDRLQSLRSERNMMNIANGPH 341
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 425 LYQKRVLTITGICVALLVVGICVWAYCKTKQKQRRQMHHLRQNMCPAHQ--RSLANGPS 483
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 342 HNPPTPENVLQNVQYKKNVISEHIVERAEETSFSTSHYTAHSTTVTQT-----PS 396
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 484 HPLRDPPEIOMAD-YISKVNPATDHYIRREAEETFSGSHSCSPSHHCSTATPTSSHRHES 542
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 397 HWSNGHTESIILSEHSIVTMSVENSRRHSPT--GGPRGRNGTGGPRECNGLRHARET 455
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 543 HTWSLERSLSLSDSGIMLSVGTCKNSCPACVEARARAAAYSQEE-----RRRAAMP 598
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 456 P-----DSYRDSPHSERYVSAMTTPARMSVDFHTPSPKSPSEMSPPVSVMTVSPSM 510
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 599 PYHDSIDSLSDSPHSERYVSALTTPARLSVDFHYSLATQVPTFEITSPNSAHAVSLPPA 658
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 511 A-VSPFMEERERPLLVTPLRLEKFDHP-----QPFSSFHENPAHDS-- 553
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 659 APISYRLAEQQPLL-----RHPAPPGPGPGGADMQRSYDSYIYYPAAAGPGPR 705
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 554 -----NSLPASPLRIVEDEYETTOEYEPQAEVKKL-----ANSRRAKRTK 595
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 706 RGACALGSLGSLPSPFRIPEDEYETTOEYECAPPPPPRPRTRGASRRTSAGPRWRRSR 765
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 596 PNGHTIANR-----LEVDNTSSSSNSESETEDEVRG-----EDTPFLGIQNPL-AASL 643
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 766 LNGLAAQRAARAARDLSLSGSGCGSASASDDDDADGALAAESTPFLGLRAAHDLRS 825
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 644 EATPAFLADSR 656
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 826 DSPPLCPLAADSR 838
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

RESULT 15
A56210
neu differentiation factor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Aug-2002
C:Accession: A56210
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;
Mol. Cell. Biol. 14, 1509-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863; PMID:7509448
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-230 <RES>

A:Cross-references: EMBL:U02315; NID:g408380; PIDN:AAAL9940.1; PID:g408381
C:Superfamily: human heregulin; EGF homology

```
Query Match      28.4%; Score 985; DB 2; Length 230;
Best Local Similarity 90.7%; Pred. No. 4.1e-48;
Matches 194; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

QY 56 KGKKKRGSGKKPESAGSQSPALPPRLKEMKQESAKLVLCETSSSEYSLRFKW 115
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 1 KGKKKDRGSRGKPGPAEGDPSPALPPRLKEMKQESAKLVLCETSSSEYSLRFKW 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 116 KGNELNRKNPKQNIQKPKGSELIRKASLADSGEYMKVSKLGNDSASANIT 175
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 61 KGNELNRKNPKQNIQKPKGSELIRKASLADSGEYMKVSKLGNDSASANIT 120
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 176 SNEITGMPASTGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTECV 235
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 121 SNEFITGMPASTETAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTECV 180
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

QY 236 NGGECFMVKDLSNPSRYLCKCPGFTGARCTENV 269
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 181 NGGECFTVKDLSNPSRYLCKCPNEFTGDRCONV 214
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
```

Search completed: July 9, 2003, 10:48:08
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:23:15 ; Search time 15 Seconds
(without alignments)
1849.844 Million cell updates/sec

Title: US-10-022-609-11

Perfect score: 3470

Sequence: 1 ARAPQGRSLSPSRDKLPN.....RLASRTNPAGRFSTQBEIQ 669

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3242	93.4	639	1 NRG1_HUMAN	Q02297 h pro-neure
2	2917	84.1	602	1 NRG1_RAT	P43322 r pro-neure
3	2314	66.7	602	1 NRG1_CHICK	Q05199 gallus gall
4	2086.5	60.1	677	1 NRG1_XENLA	O93383 xenopus lae
5	1013.5	29.2	850	1 NRG2_HUMAN	O14511 homo sapien
6	1001.5	28.9	868	1 NRG2_RAT	O35569 rattus norv
7	987.5	28.5	756	1 NRG2_MOUSE	P56974 mus musculu
8	301	8.7	713	1 NRG3_MOUSE	O35181 mus musculu
9	300	8.6	720	1 NRG3_HUMAN	P56975 homo sapien
10	273	7.9	296	1 SMDF_HUMAN	Q15491 homo sapien
11	174.5	5.0	704	1 NP14_RAT	P41777 rattus norv
12	165	4.8	1091	1 NCAL_CHICK	P13590 gallus gall
13	162	4.7	1306	1 MSB2_YEAST	P32334 saccharomyc
14	160.5	4.6	699	1 NP14_HUMAN	Q14978 homo sapien
15	157.5	4.5	3924	1 ANK2_HUMAN	Q01484 homo sapien
16	156	4.5	992	1 FRQ_LEPAU	Q01115 leptosphaer
17	154	4.4	1943	1 PC15_MOUSE	Q99pjl1 mus musculu
18	152	4.4	2843	1 APC_HUMAN	P25054 homo sapien
19	151.5	4.4	785	1 IF16_HUMAN	Q16666 homo sapien
20	150	4.3	1189	1 YJH6_YEAST	P47035 saccharomyc
21	149	4.3	906	1 CENC_MOUSE	P49452 mus musculu
22	148	4.3	633	1 MLH_TETTH	P40631 tetrahymena
23	147.5	4.3	1794	1 YAV1_SCHPO	Q10172 schizosacch
24	147	4.2	862	1 CD22_MOUSE	P35329 mus musculu
25	146.5	4.2	1092	1 NCAL_XENLA	P36335 xenopus lae
26	146	4.2	1636	1 BUD3_YEAST	P25558 saccharomyc
27	145.5	4.2	1115	1 NCAL_MOUSE	P13595 mus musculu
28	145	4.2	1411	1 TCOF_HUMAN	Q13428 homo sapien
29	143	4.1	4393	1 PGBM_HUMAN	P98160 homo sapien
30	142	4.1	3329	1 BRC2_MOUSE	P97929 mus musculu
31	140	4.0	1849	1 IGA4_HAEIN	P45386 haemophilus
32	139.5	4.0	323	1 FCGC_HUMAN	P31995 homo sapien
33	139.5	4.0	601	1 YFK5_SCHPO	P87132 schizosacch

RESULT 1	
NRG1_HUMAN	
AC Q02297; Q02298; Q02299; Q12779; Q12780; Q12781; Q12782; Q12783;	STANDARD; PRT; 639 AA.
AC Q12784; Q07110; Q07111; Q9UPE3; Q14667;	
DT 16-OCT-2001 (Rel. 40, Created)	
DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DT 15-JUN-2002 (Rel. 41, Last annotation update)	
DE Pro-neuregulin-1 precursor (Pro-NRG1) [Contains: Neuregulin-1 (Neu	
DE differentiation factor) (Heregulin) (HRG) (Breast cancer cell	
DE differentiation factor p45) (Acetylcholine receptor inducing activity)	
DE (ARIA) (Sensory and motor neuron-derived factor) (Glial growth	
DE factor)]	
GN NRG1 OR HGL OR NDF OR HRGA OR GGF OR SMDF.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA1; BETA2 AND BETA3), AND	
RP PARTIAL SEQUENCE.	
RX MEDLINE-92271253; PubMed-1350381;	
RA Holmes W.E., Sliwowski M.X., Akita R.W., Henzel W.J., Lee J.,	
RA PATAI W., Yansura D., Abadi N., Raab H., Lewis G.D., Shepard H.M.,	
RA Kuang W.-J., Wood W.I., Goeddel D.V., Vandlen R.L.;	
RL "Identification of heregulin, a specific activator of p185erbB2.";	
RN Science 256:1205-1210(1992).	
[2]	
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA1; ALPHA2B; ALPHA3; BETA1; BETA2	
RP AND BETA3).	
RC TISSUE-Pituitary, and kidney adenocarcinoma;	
RC MEDLINE-94158863; PubMed-7509448;	
RA Wen D., Suggs S.V., Karunakaran D., Liu N., Cupples R.L., Luo Y.,	
RA Janssen A.M., Ben-Baruch N., Trollinger D.B., Jacobsen V.L.,	
RA Meng S.-Y., Lu H.S., Hu S., Chang D., Yang W., Yanigahara D.,	
RA Koski R.A., Yarden Y.;	
RT "Structural and functional aspects of the multiplicity of Neu	
RT differentiation factors.";	
RL Mol. Cell. Biol. 14:1909-1919(1994).	
RN [3]	
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).	
RX MEDLINE-92208945; PubMed-1348215;	
RA Peles E., Bacus S.S., Koski R.A., Lu H.S., Wen D., Ogden S.G.,	
RA Levy R.B., Yarden Y.;	
RT "Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein	
RT that induces differentiation of mammary tumor cells.";	
RL Cell 69:205-216(1992).	
RN [4]	
RP SEQUENCE FROM N.A. (ISOFORMS BETA3 AND GGF2).	
RC TISSUE-Brain;	
RC MEDLINE-93205115; PubMed-8096067;	
RA Marchionni M.A., Goodearl A.D.J., Chen M.S., Bermingham-McDonogh O.,	
RA Kirk C., Hendricks M., Danahy F., Misumi D., Sudhalter J.,	
RA Kobayashi K., Wrblewski D., Lynch C., Baldasare M., Hiles I.,	
RA Davis J.B., Hsuan J.J., Totty N.F., Otsu M., McBurney R.N.,	
RA Waterfield M.D., Stroobant P., Gwynne D.;	

34	139.5	4.0	1133	1 EGF_RAT	P07522 rattus norv
35	139.5	4.0	2717	1 ZEP1_HUMAN	P15822 homo sapien
36	139	4.0	862	1 SMP2_YEAST	P32567 saccharomyc
37	138.5	4.0	2845	1 APC_MOUSE	Q61315 mus musculu
38	138	4.0	764	1 ICCR_DROME	Q08180 drosophila
39	137	3.9	1468	1 FMN1_MOUSE	Q05860 mus musculu
40	136.5	3.9	1906	1 KMLS_CHICK	P11799 gallus gall
41	136.5	3.9	2842	1 APC_RAT	P70478 rattus norv
42	135.5	3.9	936	1 FHL1_YEAST	P39521 saccharomyc
43	135.5	3.9	952	1 YK15_CAEEL	P46012 caenorhabdi
44	135.5	3.9	1087	1 POM1_SCHPO	Q09690 schizosacch
45	135	3.9	1163	1 RTN4_RAT	Q9Jk11 rattus norv

ALIGNMENTS

RT "Glial growth factors are alternatively spliced erbB2 ligands
 RL expressed in the nervous system.";
 RN Nature 362:312-318(1993).
 [5]
 RP SEQUENCE FROM N.A. OF GAMMA-HERGULIN FUSION PROTEIN.
 RC TISSUE-Breast cancer;
 RX MEDLINE-97472144; PubMed-9333014;
 RA Schafer G., Fitzpatrick V.D., Sliwowski M.X.;
 RT "Gamma-hergulin: a novel heregulin isoform that is an autocrine
 RL growth factor for the human breast cancer cell line, MDA-MB-175";
 RN Oncogene 15:1385-1394(1997).
 [6]
 RP SEQUENCE OF 1-210 FROM N.A.
 RA Schoumacher F., Herzer S., Flury N., Kueng W., Mueller H.,
 RL Eppenberger U.;
 RN Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 [7]
 RP SEQUENCE OF 19-27.
 RX MEDLINE-93366731; PubMed-7689552;
 RA Culouscou J.-M., Plozman G.D., Carlton G.W., Green J.M., Shoyab M.;
 RT "Characterization of a breast cancer cell differentiation factor that
 RL specifically activates the HER4/p180erbB4 receptor.";
 RN J. Biol. Chem. 268:18407-18410(1993).
 [8]
 RP CHROMOSOMAL TRANSLOCATION
 RX MEDLINE-99455251; PubMed-10523851;
 RA Wang X.-Z., Jolicoeur E.M., Conte N., Chaffanet M., Zhang Y.,
 RT Mozziconacci M.-J., Feiner H., Birnbaum D., Pebusque M.-J., Ron D.;
 RL "Gamma-hergulin is the product of a chromosomal translocation fusing
 RT the DOC4 and HGL/NRG1 genes in the MDA-MB-175 breast cancer cell
 RL line.";
 RN Oncogene 18:5718-5721(1999).
 [9]
 RP CHROMOSOMAL TRANSLOCATION
 RX MEDLINE-20065180; PubMed-10597312;
 RA Liu X., Baker E., Eyre H.J., Sutherland G.R., Zhou M.;
 RT "Gamma-hergulin: a fusion gene of DOC-4 and neuregulin-1 derived from
 RL a chromosome translocation.";
 RN Oncogene 18:7110-7114(1999).
 [10]
 RP STRUCTURE BY NMR OF 175-241 (ISOFORM ALPHA).
 RX MEDLINE-94341264; PubMed-8062828;
 RA Nagata K., Kohda D., Hatanaka H., Ichikawa S., Matsuda S.,
 RA Yamamoto T., Suzuki A., Inagaki F.;
 RT "Solution structure of the epidermal growth factor-like domain of
 RL heregulin-alpha, a ligand for p180erbB-4.";
 RN EMBO J. 13:3517-3523(1994).
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
 CC ACTIVATION OF THE ERBB RECEPTORS. THE MULTIPLE ISOFORMS PERFORM
 CC DIVERSE FUNCTIONS SUCH AS INDUCING GROWTH AND DIFFERENTIATION OF
 CC EPITHELIAL, GLIAL, NEURONAL, AND SKELETAL MUSCLE CELLS; INDUCING
 CC EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC VESICLES DURING
 CC THE FORMATION OF THE NEUROMUSCULAR JUNCTION; STIMULATING
 CC LOBULOALVEOLAR BUDDING AND MILK PRODUCTION IN THE MAMMARY GLAND
 CC AND INDUCING DIFFERENTIATION OF MAMMARY TUMOR CELLS; STIMULATING
 CC SCHWANN CELL PROLIFERATION; IMPLICATION IN THE DEVELOPMENT OF THE
 CC MYOCARDIUM SUCH AS TRABECULATION OF THE DEVELOPING HEART.
 CC -1- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN
 CC REGION OF LIMK1 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE. THE SECRETED
 CC ISOFORM GGF2 HAS A SIGNAL PEPTIDE. THE ISOFORM BETA3 MAY BE
 CC NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 10 ISOFORMS: ALPHA (SHOWN HERE);
 CC ALPHAIIA, ALPHAIIB, ALPHAIIC, BETA1, BETA2, BETA3/GGFHFB1,
 CC GGF2/GGFHFP2 AND SMDF (AC Q15491); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THEY HAVE BEEN CLASSIFIED AS TYPE I NRGS (VARIANTS WITH
 CC AN IG DOMAIN AND A GLYCOSYLATION DOMAIN; ALPHA AND BETA), TYPE II
 CC NRGS (VARIANTS WITH AN IG DOMAIN BUT NO GLYCOSYLATION DOMAIN;
 CC GGF2) AND TYPE III NRGS (VARIANTS WITH A CYS-RICH DOMAIN; SMDF).
 CC

CC ALL THESE ISOFORMS PERFORM DISTINCT TISSUE-SPECIFIC FUNCTIONS.
 CC -1- TISSUE SPECIFICITY: TYPE I ISOFORMS ARE THE PREDOMINANT FORMS
 CC EXPRESSED IN THE ENDOCARDIUM. ISOFORM ALPHA IS EXPRESSED IN
 CC BREAST, OVARY, TESTIS, PROSTATE, HEART, SKELETAL MUSCLE, LUNG,
 CC PLACENTA LIVER, KIDNEY, SALIVARY GLAND, SMALL INTESTINE AND BRAIN,
 CC BUT NOT IN UTERUS, STOMACH, PANCREAS, AND SPLEEN. ALPHAIIC IS THE
 CC PREDOMINANT FORM IN MESENCHYMAL CELLS AND IN NONNEURONAL ORGANS,
 CC WHEREAS BETA1 IS THE MAJOR NEURONAL FORM. BETA3 IS EXPRESSED IN
 CC SPINAL CORD AND BRAIN. GGF2 IS THE MAJOR FORM IN SKELETAL MUSCLE
 CC CELLS; IN THE NERVOUS SYSTEM IT IS EXPRESSED IN SPINAL CORD AND
 CC BRAIN. ALSO DETECTED IN ADULT HEART, PLACENTA, LUNG, LIVER,
 CC KIDNEY, AND PANCREAS.
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE AT EARLY EMBRYONIC AGES.
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
 CC DIMERIZATION (BY SIMILARITY).
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
 CC DOMAIN.
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
 CC FORM.
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
 CC SIMILARITY).
 CC -1- DISEASE: INVOLVED IN A RARE T(8;11) CHROMOSOMAL TRANSLOCATION THAT
 CC FUSES THE 5' END OF OD24 TO NRG1 (ISOFORM BETA3). THE PRODUCT OF
 CC THIS TRANSLOCATION WAS FIRST THOUGHT TO BE AN ALTERNATIVELY
 CC SPLICED ISOFORM, CALLED GAMMA-HERGULIN. GAMMA-HERGULIN IS A
 CC SOLUBLE ACTIVATING LIGAND FOR THE ERBB2-ERBB3 RECEPTOR COMPLEX AND
 CC ACTS AS AN AUTOCRINE GROWTH FACTOR IN A SPECIFIC BREAST CANCER
 CC CELL LINE (MDA-MB-175). NOT DETECTED IN BREAST CARCINOMA SAMPLES,
 CC INCLUDING DUCTAL, LOBULAR, MEDULLARY, AND MUCINOUS HISTOLOGICAL
 CC TYPES, NEITHER IN OTHER BREAST CANCER CELL LINES.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M94165; AAA58638.1; -
 CC EMBL; M94166; AAA58639.1; -
 CC EMBL; M94167; AAA58640.1; -
 CC EMBL; M94168; AAA58641.1; -
 CC EMBL; L12261; AAB59358.1; -
 CC EMBL; U02325; AAA19950.1; -
 CC EMBL; U02326; AAA19951.1; -
 CC EMBL; U02327; AAA19952.1; -
 CC EMBL; U02328; AAA19953.1; -
 CC EMBL; U02329; AAA19954.1; -
 CC EMBL; U02330; AAA19955.1; -
 CC EMBL; L12260; AAB59622.1; -
 CC EMBL; AF026146; RAD01795.1; -
 CC EMBL; AF009227; AAC51756.1; ALT_INIT.
 CC PDB; 1HRE; 15-OCT-94.
 CC PDB; 1HRF; 15-OCT-94.
 CC Genew; HGNC:7997; NRG1.
 CC MIM; 142445; -
 CC InterPro; IPR003006; Iq.MHC.
 CC InterPro; IPR002154; Neuregulin.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00047; Iq; 1.
 CC Pfam; PF02158; Neuregulin; 1.
 CC PRINTS; PR01089; NEUREGULIN.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; 1.
 CC Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
 CC

Query Match	93.48;	Score	3242;	DB	1;	Length	639;
Best Local Similarity	100.0%;	Pred.	No. 5.6e-173;				
Matches	624;	Conservative	0;	Mismatches	0;	Indels	0;
							Gaps
QY	46	SERKEGKGKGGKKKGGKPPSAAGSOSPAPPRLLKEMKSOESAAGSKLVLRCTSS	105				
DB	1	SERKEGKGKGGKKKGGKPPSAAGSOSPAPPRLLKEMKSOESAAGSKLVLRCTSS	60				
QY	106	EYSLRFPKFNKGNELNRKNKPQNIQKPKGKSELRIKASLADSGEYCMKVISKLND	165				
DB	61	EYSLRFPKFNKGNELNRKNKPQNIQKPKGKSELRIKASLADSGEYCMKVISKLND	120				
QY	166	SASANTIVESNELIITGMPASTEGAYVSSSPPIRISVSTEGANTSSSTSTTTGTSHLVK	225				
DB	121	SASANTIVESNELIITGMPASTEGAYVSSSPPIRISVSTEGANTSSSTSTTTGTSHLVK	180				
QY	226	CAKEKFTFCVNGGBCFMVKDLNPNRSLVLCQPGFTGARCENPVMKVONQEKABEELYOK	285				
DB	181	CAKEKFTFCVNGGBCFMVKDLNPNRSLVLCQPGFTGARCENPVMKVONQEKABEELYOK	240				
QY	286	RVLTITGICITALLVVGIMCVVAYCKTKKQKKLHDLRQSLRSERNMMNTANGPHHPNP	345				
DB	241	RVLTITGICITALLVVGIMCVVAYCKTKKQKKLHDLRQSLRSERNMMNTANGPHHPNP	300				
QY	346	PPENVQLVNOYVSKNVLSSHIVREARETFSFSTSHYTSTAHHSTVTQTTPSHSWSNGHTE	405				
DB	301	PPENVQLVNOYVSKNVLSSHIVREARETFSFSTSHYTSTAHHSTVTQTTPSHSWSNGHTE	360				
QY	406	SILSESHSVIVMSSVENSRRSSPPGPRGLNGTGGPRECNFSFURHARETSDSYRDSPHS	465				
DB	361	SILSESHSVIVMSSVENSRRSSPPGPRGLNGTGGPRECNFSFURHARETSDSYRDSPHS	420				
QY	466	ERYVSAMTTTPARKSPVDFTHTPSSPKPPSEMSPPVSSMTVSMPSMAYSPFMEERPLILV	525				
DB	421	ERYVSAMTTTPARKSPVDFTHTPSSPKPPSEMSPPVSSMTVSMPSMAYSPFMEERPLILV	480				
QY	526	TPPRLRKKFDHHPQQFSFPHHNPADNSLPA SPLRIVEDEEYETQETEPQAPQVKKL	585				
DB	481	TPPRLRKKFDHHPQQFSFPHHNPADNSLPA SPLRIVEDEEYETQETEPQAPQVKKL	540				
QY	586	ANSRRAKRTKPNGHIANRLEVDNSTSSQSNSESETDERVGEDTFFLGIONPLAASLEA	645				
DB	541	ANSRRAKRTKPNGHIANRLEVDNSTSSQSNSESETDERVGEDTFFLGIONPLAASLEA	600				
QY	646	TPAFRLADSRTPAGRFSTQEEIQ	669				
DB	601	TPAFRLADSRTPAGRFSTQEEIQ	624				

RESULT 2

[illegible]

Db 351 TSFSTSHVTSHTSHSTVTTQTPSHSNGLSEMSISEKSYSVIVTSSVENSHTSPT-GP 409
 QY 433 RGRNGTGPRCNSFLRHARETDPDSYRSPHSERYVSAMTPARMSPVDFHTPPSPKSP 492
 Db 410 RGRNGIGGPRDC-SYLHARDTPDSYRSPHSERYVSAMTPARMSPVDFHTPPSPKSP 468
 QY 493 PSEMPVPVSTVMSVNAVSPFMEERPLLVTPPRLREKFDH-PQ-----QFSSEH 546
 Db 469 CLETPSPSSSLAVSPVAVSFIEERPLLVSPRLREKRYDRKTPQKPHKQHSYH 528
 QY 547 HNPADHSNLSPLRIVEDEYETQYEPAQEPVKKLANRRRAKTKPNCHIANRLEV 606
 Db 529 HNPGDSSSLPNLRIVEDEYETQYEPSPLEPAKLVNSRRKTKPNCHISNRLEL 588
 QY 607 DNTSSQSSNSSETEDETVGDDTFFLQNPPLAASLEATPAFLADSRTPNAGRFSTQE 666
 Db 589 DSDSSSESTSETEDETVGDDTFFLQNPPLAASLEASLYRADSRTPNAGRFSTQE 648
 QY 667 EQQ 669
 649 EQQ 651

RESULT 5

NRG2_HUMAN
 ID NRG2_HUMAN STANDARD; PRT; 850 AA.
 AC O14511;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pro-neuregulin-2 precursor (Pro-NRG2) [Contains: Neuregulin-2 (NRG-2)
 DE (Neural- and thymus-derived activator for ERBB kinases) (NTAK)
 DE (Divergent of neuregulin 1) (DON-1)].
 GN NRG2 OR NTAK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Neuroblastoma;
 RX MEDLINE=9806324; PubMed=9348101;
 RA Higashiyama S., Horikawa M., Yamada K., Ichino N., Nakano N.,
 RA Nakagawa T., Miyagawa J., Matsushita N., Nagatsu T., Taniguchi N.,
 RA Ishiguro H.;
 RT "A novel brain-derived member of the epidermal growth factor family
 RT that interacts with ErbB3 and ErbB4.";
 J. Biochem. 122:675-680(1997).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORMS DON-1B AND DON-1R).
 RC TISSUE=Fetal brain;
 RX MEDLINE=97342638; PubMed=9199335;
 RA Busfield S.J., Michnick D.A., Chickering T.W., Revett T.L., Ma J.,
 RA Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Goodearl A.D.J.,
 RA Gearing D.P.;
 RT "Characterization of a neuregulin-related gene, Don-1, that is highly
 RT expressed in restricted regions of the cerebellum and hippocampus.";
 RL Mol. Cell. Biol. 17:4007-4014(1997).
 [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6).
 RC TISSUE=Lung, and Fetal brain;
 RX MEDLINE=99295836; PubMed=10369162;
 RA Ring H.Z., Chang H., Guilbot A., Brice A., LeGuern E., Francke U.;
 RT "The human neuregulin 2 (NRG2) gene: cloning, mapping and evaluation
 RT as a candidate for the autosomal recessive form of Charcot-Marie-Tooth
 RT disease linked to 5q.";
 RL Hum. Genet. 104:326-332(1999).
 [3]
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
 CC ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE
 CC HETERODIMERIZATION WITH THE EGF RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS

A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
 MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 8 ISOFORMS; 1 (SHOWN HERE), 2, 3, 4, 5, 6,
 CC DON-1B AND DON-1R; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE CEREBELLUM IN THE ADULT.
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
 CC DIMERIZATION (BY SIMILARITY).
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
 CC DOMAIN (BY SIMILARITY).
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
 CC FORM (BY SIMILARITY).
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; AB005060; BAA23417.1;
 CC EMBL; AF119162; AAF28848.1;
 CC EMBL; AF119151; AAF28848.1; JOINED.
 CC EMBL; AF119152; AAF28848.1; JOINED.
 CC EMBL; AF119153; AAF28848.1; JOINED.
 CC EMBL; AF119154; AAF28848.1; JOINED.
 CC EMBL; AF119155; AAF28848.1; JOINED.
 CC EMBL; AF119156; AAF28848.1; JOINED.
 CC EMBL; AF119157; AAF28848.1; JOINED.
 CC EMBL; AF119158; AAF28848.1; JOINED.
 CC EMBL; AF119159; AAF28848.1; JOINED.
 CC EMBL; AF119160; AAF28848.1; JOINED.
 CC EMBL; AF119161; AAF28848.1; JOINED.
 CC EMBL; AF119162; AAF28848.1; JOINED.
 CC EMBL; AF119163; AAF28848.1; JOINED.
 CC EMBL; AF119164; AAF28848.1; JOINED.
 CC EMBL; AF119165; AAF28848.1; JOINED.
 CC EMBL; AF119166; AAF28848.1; JOINED.
 CC EMBL; AF119167; AAF28848.1; JOINED.
 CC EMBL; AF119168; AAF28848.1; JOINED.
 CC EMBL; AF119169; AAF28848.1; JOINED.
 CC EMBL; AF119170; AAF28848.1; JOINED.
 CC EMBL; AF119171; AAF28848.1; JOINED.
 CC EMBL; AF119172; AAF28848.1; JOINED.
 CC EMBL; AF119173; AAF28848.1; JOINED.
 CC EMBL; AF119174; AAF28848.1; JOINED.
 CC EMBL; AF119175; AAF28848.1; JOINED.
 CC EMBL; AF119176; AAF28848.1; JOINED.
 CC EMBL; AF119177; AAF28848.1; JOINED.
 CC EMBL; AF119178; AAF28848.1; JOINED.
 CC EMBL; AF119179; AAF28848.1; JOINED.
 CC EMBL; AF119180; AAF28848.1; JOINED.
 CC EMBL; AF119181; AAF28848.1; JOINED.
 CC EMBL; AF119182; AAF28848.1; JOINED.
 CC EMBL; AF119183; AAF28848.1; JOINED.
 CC EMBL; AF119184; AAF28848.1; JOINED.
 CC EMBL; AF119185; AAF28848.1; JOINED.
 CC EMBL; AF119186; AAF28848.1; JOINED.
 CC EMBL; AF119187; AAF28848.1; JOINED.
 CC EMBL; AF119188; AAF28848.1; JOINED.
 CC EMBL; AF119189; AAF28848.1; JOINED.
 CC EMBL; AF119190; AAF28848.1; JOINED.
 CC EMBL; AF119191; AAF28848.1; JOINED.
 CC EMBL; AF119192; AAF28848.1; JOINED.
 CC EMBL; AF119193; AAF28848.1; JOINED.
 CC EMBL; AF119194; AAF28848.1; JOINED.
 CC EMBL; AF119195; AAF28848.1; JOINED.
 CC EMBL; AF119196; AAF28848.1; JOINED.
 CC EMBL; AF119197; AAF28848.1; JOINED.
 CC EMBL; AF119198; AAF28848.1; JOINED.
 CC EMBL; AF119199; AAF28848.1; JOINED.
 CC EMBL; AF120000; AAF28848.1; JOINED.

RA Ishiguro H.;
 RT "A novel brain-derived member of the epidermal growth factor family
 RL that interacts with ErbB3 and ErbB4.";
 RN J. Biochem. 122:675-680(1997).
 [2]
 RC SEQUENCE OF 109-868 FROM N.A. (ISOFORMS NR2-ALPHA AND NR2-BETA).
 RD TISSUE=Cerebellum;
 RE MEDLINE=973111397; PubMed=9168114;
 RF Chang H., Riese D.J. II, Gilbert W., Stern D.F., McMahon U.J.;
 RG "Ligands for ErbB-family receptors encoded by a neuregulin-like
 RT gene.";
 CC Nature 387:509-512(1997).
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORRECEPTORS,
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
 CC ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE
 CC HETERODIMERIZATION WITH THE EGF RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 7 ISOFORMS; NTA-ALPHA (SHOWN
 CC HERE), NTA-ALPHA2A, NTA-ALPHA2B/NTAK-ALPHA2-1P, NTA-ALPHA2, NTA-
 CC GAMMA, NR2-ALPHA AND NR2-BETA; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THE ALPHA-TYPE AND BETA-TYPE DIFFER IN THE EGF-LIKE
 CC DOMAIN.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST PARTS OF THE BRAIN,
 CC ESPECIALLY THE OLFACTORY BULB AND CEREBELLUM WHERE IT LOCALIZES IN
 CC GRANULE AND PURKINJE CELLS. IN THE HIPPOCAMPUS, FOUND IN THE
 CC GRANULE CELLS OF THE DENTATE GYRUS. IN THE BASAL FOREBRAIN, FOUND
 CC IN THE CHOLINERGIC CELLS. IN THE HINDRAIN, WEAKLY DETECTABLE IN
 CC THE MOTOR TRIGEMINAL NUCLEUS. NOT DETECTED IN THE HYPOTHALAMUS.
 CC ALSO FOUND IN THE LIVER AND IN THE THYMUS. NOT DETECTED IN HEART,
 CC ADRENAL GLAND, OR TESTIS.
 CC -1- DEVELOPMENTAL STAGE: IN THE EMBRYO, EXPRESSED IN THE BRAIN OF
 CC E11.5 EMBRYOS WHERE IT IS FOUND IN THE TELECEPHALON, BUT NOT IN
 CC THE HINDBRAIN. NOT FOUND IN THE HEART. IN THE ADULT, FOUND IN
 CC BRAIN AND THYMUS
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
 CC DIMERIZATION (BY SIMILARITY).
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
 CC DOMAIN (BY SIMILARITY).
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
 CC FORM (BY SIMILARITY).
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D89995; BAA23344.1; -
 CC EMBL: D89996; BAA23345.1; -
 CC EMBL: D89997; BAA23346.1; -
 CC EMBL: D89998; BAA23347.1; -
 CC EMBL: AB001576; BAA23348.1; -
 CC HSP: Q12784; 1HRE.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR003006; Ig_MHC.
 CC InterPro: IPR003598; Ig_c2.
 CC InterPro: IPR001594; Neuregulin.
 CC Pfam: PF00008; EGF; 1.
 CC Pfam: PF00047; Ig; 1.
 CC Pfam: PF02158; Neuregulin; 1.

DR SMART: SM00181; EGF; 1.
 DR SMART: SM00408; Igc2; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 KW Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
 KW Transmembrane; Multigene family; Alternative splicing.
 FT PROPEP 1 127
 FT CHAIN 128 868 PRO-NEUREGULIN-2, MEMBRANE-BOUND FORM.
 FT CHAIN 128 868 NEUREGULIN-2.
 FT DOMAIN 128 428 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 128 429 INTERNAL SIGNAL SEQUENCE (POTENTIAL).
 FT TRANSMEM 430 450 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 451 868 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 266 334 SER/THR-RICH.
 FT DOMAIN 346 356 EGF-LIKE.
 FT DOMAIN 357 398 POLY-SER.
 FT DOMAIN 22 32 POLY-THR.
 FT DOMAIN 35 45 POLY-ALA.
 FT DOMAIN 56 59 POLY-PRO.
 FT DOMAIN 103 106 BY SIMILARITY.
 FT DISULFID 273 327 BY SIMILARITY.
 FT DISULFID 361 375 BY SIMILARITY.
 FT DISULFID 369 386 BY SIMILARITY.
 FT DISULFID 388 397 BY SIMILARITY.
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 362 362 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1 108 MISSING (IN ISOFORM NR2-BETA).
 FT VARSPLIC 220 222 PLV -> FFF (IN ISOFORM NTA-ALPHA2-1P).
 FT VARSPLIC 388 388 MISSING (IN ISOFORM NTA-ALPHA2-1P).
 FT VARSPLIC 389 868 NGFFGRCLEKPLRLYMPDPKQ -> VGYTDRCOQFAMV
 FT VARSPLIC 390 412 NFA (IN ISOFORM NR2-BETA).
 FT VARSPLIC 390 421 NGFFGRCLEKPLRLYMPDPKQKHLGFELKE -> VGYTDRCOQFAMV
 FT VARSPLIC 414 421 MISSING (IN ISOFORM NTA-ALPHA2A AND
 FT VARSPLIC 414 439 ISOFORM NTA-ALPHA2B).
 FT VARSPLIC 440 868 HLGFEKLEKPLRLYMPDPKQKHLGFELKE -> SVLWDPGPGV
 FT CONFLICT 117 117 S -> F (IN REF. 2).
 FT CONFLICT 724 724 S -> H (IN REF. 2).
 SQ SEQUENCE 868 AA; 93776 MW; 3C7D4D94DBE64DE2 CRC64;
 Query Match 28.9%; Score 1001.5; DB 1; Length 868;
 Best Local Similarity 36.8%; Pred. No. 1.6e-48;
 Matches 248; Conservative 102; Mismatches 184; Indels 139; Gaps 21;
 QY 55 GKGKKKRGSGKKPSAAGSQSPALPPRLKMKSGESAAGSKVLRLCEFTSSYSLRPFKW 114
 DB 234 GKNIRKEVGKILTCAT-----RPKLKMKSGTGEVGEKQSLKCEAAAGNPPQPSYRW 286
 QY 115 FRNGELNRKPNQIKIQQKPKG--SELIRKASLADSGEYMKVSKLGNDSASANT 172
 DB 287 FKDGKELNRS---RDRIKYGKRNKSRLOFNKVKVEDAGEYVCEAENLTKDVTYGRHL 343
 QY 173 IVESNEIITGMPASTGAYVSSPIRISVSTEGANTSSSTSTSTSTSTSHLVKCAEKET 232
 DB 344 V-----NSVSTTSSWSG--HARKCNETAKS 367
 QY 233 FCVNGGECFPMVKDLNPNRSLKCKCPGFTGARTENPMKVQ-----NOEKABE 281
 DB 368 YCVNGGVYIEGINKLS---CKCPNGPFGQCKLEKPLRLYMPDPKQKHLGFELKEABE 424
 QY 282 LYQKRVLTITGICIAALLVVGIMCVAYCKTKKQKRLHDLRQLRSERNRNMANGPH 341
 DB 425 LYQKRVLTITGICVALLVVGIMCVAYCKTKKQKRLHDLRQLRSERNRNMANGPH 483
 QY 342 HNPPEPVQVNOVYKKNVISSEHIVERAEATSTSTSHYTAHSTTVTQT-----PS 396
 DB 484 HPRDPEIQMAD--YISKNPVPAIDHVRIRREATSTSGSHSCSPSHHCSTATPTSSHRHES 542

QY 397 HSWNGHTESILSHSVIVSSVNSRHSPT-GGPRGLNCTGGPRECSFLRHARET 455
 DB 543 HTWSERSESLTDSQSGIMLSSVGTSCNCPACVEARAAAYSQEE-----RRRAAMP 598
 QY 456 P-----DSYRDSHSEYVSAMTTPARMSVDPHTSPSPKSPSEMSPPVSSMTVSPSM 510
 DB 599 PYHDSIDSLDHSRSEYVSALTTPARLSVDPVDFHYSLATQVPTFEITSPNSAHAVSLPPA 658
 QY 511 A-VSPFMEERPLLVTPRLREKKFQHP-----QQFSSPHNPAHDS-- 553
 DB 659 APISYRLAEQPLL-----RHPAPGPGPGADMQRSYDSYYPAGGPR 705
 QY 554 -----NSLPASPLRIVEDEYETQYEPAGEPVKKL-----ANSRRRAKRTK 595
 DB 706 RGACALGSLGSLPASFRIPEDYEYETQECAPPPTPRTRGASRRTSAGPRWRRSR 765
 QY 596 PNGHIANR-----LEVDNTSSQSSNSESEYEDERVG-----EDTPFLGIONPL-AASL 643
 DB 766 LNLGAQRAARANDSLSSGCGSASASDDADDADGALAAESTPFLGLRAAHADALS 825
 644 EATPAFLADSR 656
 DB 826 DSPPLCPADSR 838

RESULT 7
 NRG2_MOUSE STANDARD; PRT; 756 AA.
 AC P56974;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pro-neuregulin-2 precursor (Pro-NRG2) [Contains: Neuregulin-2 (NRG-2)
 (Divergent of neuregulin 1) (DON-1)]
 GN NRG2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORMS NRG2-5; NRG2-10 AND NRG2-16A).
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=97342638; PubMed=9168115;
 RA Carraway K.B., Iii, Weber J.L., Unger M.J., Ledesma J., Yu N.,
 RA Gassmann M., Lai C.;
 RT "Neuregulin-2, a new ligand of ErbB3/ErbB4-receptor tyrosine
 RT kinases";
 BU Nature 387:512-516(1997).
 [2]
 RC SEQUENCE OF 150-756 FROM N.A. (ISOFORMS DON-1M AND DON-1S).
 RC TISSUE=Choroid plexus;
 RX MEDLINE=97342638; PubMed=9199335;
 RA Busfield S.J., Michnick D.A., Chickering T.W., Revett T.L., Ma J.,
 RA Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Goodearl A.D.J.,
 RA Gearing D.P.;
 RT "Characterization of a neuregulin-related gene, Don-1, that is highly
 RT expressed in restricted regions of the cerebellum and hippocampus";
 RL Mol. Cell. Biol. 17:4007-4014(1997).
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORCEPTORS,
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
 CC ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE
 CC HETERODIMERIZATION WITH THE EGF RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS; DON-1M, DON-1S/NRG2-5,
 CC NRG2-10 AND NRG2-16A (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN THE BRAIN, WITH LOWER
 CC LEVELS IN THE LUNG. IN THE CEREBELLUM, FOUND IN GRANULE AND
 CC PURKINJE CELLS.

CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
 CC DIMERIZATION (BY SIMILARITY).
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
 CC DOMAIN (BY SIMILARITY).
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
 CC FORM (BY SIMILARITY).
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
 CC HSP; Q12784; IHRF.
 DR MGD; MGI:1098246; Nrg2.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR002154; Neuregulin.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF02158; Neuregulin; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00408; IGC2; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
 KW Transmembrane; Multigene family; Alternative splicing.
 FT PROPEP 1
 FT CHAIN 20 756
 FT CHAIN 20 314
 FT DOMAIN 20 315
 FT TRANSMEM 316 336
 FT DOMAIN 337 756
 FT DOMAIN 158 226
 FT DOMAIN 238 248
 FT DOMAIN 249 290
 FT DOMAIN 627 633
 FT DISULFID 165 219
 FT DISULFID 253 267
 FT DISULFID 261 278
 FT DISULFID 280 289
 FT CARBOHYD 55 55
 FT CARBOHYD 186 186
 FT CARBOHYD 254 254
 FT CARBOHYD 296 296
 FT VARSPLIC 280 280
 FT VARSPLIC 281 756
 FT VARSPLIC 282 330
 FT
 FT
 FT VARSPLIC 331 756
 FT VARSPLIC 282 307
 FT
 FT
 SQ SEQUENCE 756 AA; 82213 MW; 51D85DC918BE678E CRC64;
 Query Match 28.5%; Score 987.5; DB 1; Length 756;
 Best Local Similarity 36.8%; Pred. No. 8.3e-48;
 Matches 246; Conservative 101; Mismatches 187; Indels 135; Gaps 21;
 QY 55 GKGKKRGSGKKPESAAQSQSPALPRLEKMSQSAAGSKLYLRCETSSYSSLRFKW 114
 DB 126 GKNKKKEVGKILCTDCAT-----RPLKKMKRSQTGEVGEKQSLKCEAAAGNPQPSYRW 178
 QY 115 FKNGNELNKRKNPQNIQKKPKCK--SELIRNKASLADSGEYCKVYSLKGNDSANIT 172
 DB 179 FDKGKELNRS---RDRIKRYGNKNSRLOFNQVRVEDAGEYVCEAENTLGDYVGRHL 235
 QY 173 IVESNEIITGMPASTEGAYVSSSESPRISVSTEGANTSSSTSTGTGTHLVKCAKEKT 232
 DB 236 V-----NSVSTLLSSWSG--HARKCNETAKS 259

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:37:10 ; Search time 88 Seconds
(without alignments)
1566.426 Million cell updates/sec

Title: US-10-022-609-11

Perfect score: 3470

Sequence: 1 ARAPQRGRSLSPSRDKLPN.....RLADSRNPAGRFSTQEEIQ 619

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

otal number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_undefined.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2800	80.7	782	11 Q9ESA5	Q9esa5 rattus norv
2	2248.5	64.8	695	11 Q9EB0	Q9esb0 rattus norv
3	2141	61.7	700	11 Q9ESB1	Q9esb1 rattus norv
4	2026.5	58.4	461	11 Q35947	Q35947 mesocricetu
5	1282.5	37.0	298	11 Q9ESA9	Q9esa9 rattus norv
6	1024	29.5	241	6 Q07112	Q07112 bos taurus
7	935.5	27.0	317	11 Q9ESA3	Q9esa3 rattus norv
8	929	26.8	342	11 Q9ESA1	Q9esa7 rattus norv
9	921	26.5	323	11 Q9ESA2	Q9esa2 rattus norv
10	273	7.9	296	4 Q96IB3	Q96ib3 homo sapien
11	269.5	7.8	111	11 Q9ESA8	Q9esa8 rattus norv
12	269	7.8	256	11 Q9ESA6	Q9esa6 rattus norv
13	266	7.7	136	11 Q9ESA7	Q9esa7 rattus norv
14	192	5.5	1514	5 Q8SY55	Q8sy55 drosophila
15	189.5	5.5	1324	5 Q8T2N1	Q8tzn1 dictyosteli
16	188.5	5.4	1106	4 Q8WX93	Q8wx93 homo sapien

17	188.5	5.4	5476	5 Q9NJ17	Q9nj17 drosophila
18	188.5	5.4	5533	5 Q9VPL2	Q9vpl2 drosophila
19	188.5	5.4	5533	5 Q9VPL3	Q9vpl3 drosophila
20	188.5	5.4	5554	5 Q9NHN1	Q9nhn1 drosophila
21	188.5	5.4	5560	5 Q9VPL1	Q9vpl1 drosophila
22	187.5	5.4	1565	5 Q8TIM2	Q8tim2 dictyosteli
23	187	5.4	3257	5 Q9V736	Q9v736 drosophila
24	180.5	5.2	5327	5 Q76891	Q76891 drosophila
25	180	5.2	897	11 Q70495	Q70495 mus musculus
26	179.5	5.2	2254	10 Q9LN02	Q9ln02 arabidopsis
27	177	5.1	2006	5 Q9WFE2	Q9wfe2 drosophila
28	172	5.0	3507	5 Q23587	Q23587 caenorhabdi
29	171	4.9	3111	5 Q9VHI0	Q9vhi0 drosophila
30	170	4.9	990	13 Q91803	Q91803 xenopus lae
31	166.5	4.8	1172	5 Q9VE53	Q9ve53 drosophila
32	166.5	4.8	1262	4 Q9UQ40	Q9uq40 homo sapien
33	166.5	4.8	1459	5 Q17084	Q17084 caenorhabdi
34	166.5	4.8	1464	5 Q61802	Q61802 caenorhabdi
35	166.5	4.8	2296	4 Q9UHA8	Q9uha8 homo sapien
36	166.5	4.8	2752	4 Q9UQ35	Q9uq35 homo sapien
37	165.5	4.8	700	4 Q9BUV3	Q9buv3 homo sapien
38	165.5	4.8	1164	5 Q9VI61	Q9vi61 drosophila
39	165	4.8	1228	11 Q9JKB5	Q9jkb5 rattus norv
40	165	4.8	1529	5 Q9GQC2	Q9gqc2 dictyosteli
41	164.5	4.7	1031	13 Q90VM2	Q90vm2 brachydanio
42	164.5	4.7	1097	4 Q8TDW7	Q8tdw7 homo sapien
43	162.5	4.7	6632	5 Q17362	Q17362 caenorhabdi
44	162.5	4.7	6632	5 Q01761	Q01761 caenorhabdi
45	161	4.6	1492	5 Q8SSU1	Q8ssul dictyosteli

ALIGNMENTS

RESULT 1

Q9ESA5 PRELIMINARY; PRT; 782 AA.
ID Q9ESA5
AC Q9ESA5;
DT 01-MAR-2001 (TREMREL. 16, Created)
DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Glial growth factor beta la (Fragment).
GN NRGL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SPINAL CORD/BRAIN STEM;
RA Carroll S.L., Stoneypher M.S., Anderson K.D., Pearson R.J. Jr.,
RA Frohnert P.W.;
RT "Structural and Functional Diversity of Glial Growth Factor Isoforms
Expressed in Regenerating Peripheral Nerve and Associated Neurons.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF194993; AAG28433.1;
DR HSSP; Q12784; 1HRE.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002114; Hpr_Serp_site.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR002154; Neuregulin.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF02158; Neuregulin; 1.
DR PRINTS; PR01089; NEUREGULIN.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.

```
KW Immunoglobulin domain.
FT NON_TER 1
SQ SEQUENCE 782 AA; 86036 MW; F6174A68F4E27BDE CRC64;

Query Match      80.7%; Score 2800; DB 11; Length 782;
Best Local Similarity 78.9%; Pred. No. 1.6e-184;
Matches 558; Conservative 31; Mismatches 58; Indels 60; Gaps 7;

QY 12 PSRDKLFPNIRALGPSNAPAPRVVER-----SVSGEMSERKEGRGK-----GK 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72 PSTDQ-----PGDPAPYLKVHVQVWAVKAGGLKNDLTLVRLDTWGHGPAFSCGR 121

QY 57 GKKEK-----GSGKKPSAAGSQSP-----ALPPLKEMK 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 LKEDSYIFFMEPDANSSGAPPAFRASFPFLETGRNLKKEVSRVLCRCALPRLKEMK 181

QY 88 SQSAAGSKVLVLCETSSSEYSSLRFKWFKNGNLNKNPNKONIKQKPKGSLRLNKAS 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 SQSAAGSKVLVLCETSSSEYSSLRFKWFKNGNLNKNPNKONIKQKPKGSLRLNKAS 241

148 LADSGEYMKVSKLGNDSASANITIVESNEITGMPASTEGAYVSSEPIRISVSTEGA 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 LADSGEYMKVSKLGNDSASANITIVESNEITGMPASTETAYVSSEPIRISVSTEGA 301

QY 208 NTSSTSTSTGTHSHLVKCAEKEKTCVNGGECFVYKDLNPNRYLCKQPGTGARCTE 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 NTSSTSTSTGTHSHLVKCAEKEKTCVNGGECFVYKDLNPNRYLCKQPGTGARCTE 361

QY 268 NVPKMYQVQ-----EKAELYQKRVLTITGICALLVVGIMCVAYCKTKKQKRLDRL 322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 YVNASPYKHLGIEFMAEELYQKRVLTITGICALLVVGIMCVAYCKTKKQKRLDRL 421

QY 323 QSLRSERNMNIANGPHHPNPPENNVQVYVSKNVISSSHIYERAEATSFSTSHYT 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 QSLRSERNMNIANGPHHPNPPENNVQVYVSKNVISSSHIYERAEATSFSTSHYT 481

QY 383 STAHSTTTVQTPSHWSNGHTESVISESHVIVSMSSVNSRHSSPTGPRGLNGTGP 442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 STAHSTTTVQTPSHWSNGHTESVISESHVIVSMSSVNSRHSSPTGPRGLNGTGP 541

QY 443 RECNSFLRHARETPDSYRSPHSERIVSAMTPARMSPVDFHTPSSPKPPSPSPVSS 502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 RD-NSFLRHARETPDSYRSPHSERIVSAMTPARMSPVDFHTPSSPKPPSPSPVSS 600

QY 503 MIVSMPSMAVSPMEERPLLLVTPRLREKFDHHPQFSSFFHHPAHDSNLSPLR 562
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 MIVSMPSMAVSPMEERPLLLVTPRLREKFDHHPQFSSFFHHPAHDSNLSPLR 660

QY 563 IVEDEEYETQVEYPAQEPVKLANSRRAKRTKPNGHIANRLEVDNSTSSQSNSESETE 622
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 IVEDEEYETQVEYPAQEPVKLANSRRAKRTKPNGHIANRLEVDNSTSSQSNSESETE 720

QY 623 DERVGEDTDFLGIQNPALASLEATPAFLADSRNTPAGRFSTQEEIQ 669
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 DERVGEDTDFLGIQNPALASLEATPAFLADSRNTPAGRFSTQEEIQ 767

RESULT 2
Q9ESB0 PRELIMINARY; PRT; 695 AA.
AC Q9ESB0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE SMDF neuregulin alpha 2a.
GN NRGI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BDIX;
```

```
RA Carroll S.L., Anderson K.D., Frohnert P.W.;
RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF194439; AG28428.1; -
DR HSSP; Q12784; IHRE.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002114; HPR_Serp_site.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF02158; Neuregulin; 1.
DR PRINTS; PRO1089; NEUREGULIN.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00001; EGF_like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
SQ SEQUENCE 695 AA; 75646 MW; 5277F2CBA2FB6878 CRC64;

Query Match      64.8%; Score 2248.5; DB 11; Length 695;
Best Local Similarity 80.4%; Pred. No. 1.3e-146;
Matches 439; Conservative 30; Mismatches 48; Indels 29; Gaps 6;

QY 151 SGEYMKVSKLGNDSASANITIVESNEITGMPAS--TEGAYVSS-----ESP 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 SSEAYTSPYVK-AQSEAGARVTVQGDHAAVASSEPSAVPTRKNRLSAFPFHTAPPPSP 195

QY 198 IRI-SYSTEGANTSS-----STSTSTGTHSHLVKCAEKEKTCVNGGECFVY 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 ARTPEVTRPKGTQPTTETNLTAPKLTSTSTGTHSHLVKCAEKEKTCVNGGECFV 255

QY 244 KDLNPNRYLCKQPGTGARCTENYPMKVQVQKAEELYQKRVLTITGICALLVVGIM 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 KDLNPNRYLCKQPGTGARCTENYPMKVQVQKAEELYQKRVLTITGICALLVVGIM 315

QY 304 CVVAYCKTKKQKRLDRLRQSLRSERNMNIANGPHHPNPPENNVQVYVSKNVI 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 CVVAYCKTKKQKRLDRLRQSLRSERNMNIANGPHHPNPPENNVQVYVSKNVI 375

QY 364 SEHIVERAEATSFSTSHYTSTAHSSTVTTQTPSHWSNGHTESVISESHVIVSMSSVENS 423
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 SEHIVERAEATSFSTSHYTSTAHSSTVTTQTPSHWSNGHTESVISESHVIVSMSSVENS 435

QY 424 RHSSPTGPRGLNGTGGPRECNSFLRHARETPDSYRSPHSERIVSAMTPARMSPVDF 483
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 RHSSPTGPRGLNGTGGPRECNSFLRHARETPDSYRSPHSERIVSAMTPARMSPVDF 494

QY 484 HTPSSPKSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 495 HTPSSPKSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 554

QY 544 SFHHPAHDSNLSPLRIVEDEEYETQVEYPAQEPVKLANSRRAKRTKPNGHIANR 603
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 555 SFHHPAHDSNLSPLRIVEDEEYETQVEYPAQEPVKLANSRRAKRTKPNGHIANR 614

QY 604 LEVDNSTSSQSNSESETEDEVRGEDTDFLGIQNPALASLEATPAFLADSRNTPAGRF 663
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 615 LEVDNSTSSQSNSESETEDEVRGEDTDFLGIQNPALASLEATPAFLADSRNTPAGRF 674

QY 664 TQEEIQ 669
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 675 TQEEIQ 680

RESULT 3
Q9ESB1 PRELIMINARY; PRT; 700 AA.
ID Q9ESB1
AC Q9ESB1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE SMDF neuregulin beta 1a.
GN NRGI.
```


DR	Pfam; PF00008; EGF; 1.
DR	Pfam; PF02158; Neuregulin; 1.
DR	PRINTS; PRO1089; NEUREGULIN.
DR	SMART; SM00181; EGF; 1.
DR	SMART; SM00001; EGF_like; 1.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
FT	NON_TER 1
FT	NON_TER 298
FT	SEQUENCE 298 AA; 32851 MW; BD76F014C2B33026 CRC64;
QY	Query Match 37.0%; Score 1282.5; DB 11; Length 298;
DB	Best Local Similarity 93.0%; Pred. No. 1.4e-80;
DB	Matches 239; Conservative 12; Mismatches 5; Indels 1; Gaps 1;
QY	212 STSTSTTGTSHLVKCAEKETFCVNGGECFMVKDLSNPRLYLCKCQPGFTGARTENVP 271
DB	: : : : : : : : :
DB	38 STSTSTTGTSHLIKAEKETFCVNGGECTFKDLSNPSRYLCKCQPGFTGARTENVP 97
QY	272 KVNQERAAELYQKRVTITIGICITALVLVGIMCVAYCKTKQSKKLHDLRQLSRSEN 331
DB	: : : : : : : :
DB	98 KVQTERAAELYQKRVTITIGICITALVLVGIMCVAYCKTKKQKQLHDLRQLSRSEN 157
QY	332 NMNIANGPHHPNPPEPNVOLNVQYSKNVISSEHIIVEREAETSFSHTYSTAHHSSTV 391
DB	: : : : : : : :
DB	158 NLVNIANGPHHPNPPEPNVOLNVQYSKNVISSEHIIVEREVETSFSTSHYTSTAHHSTV 217
QY	392 TQPPSHWSNGHTESILSEHSVIVMSVENSRSSSTGGPGRGLNGTGCGPRENSFLRH 451
DB	: : : : : : : :
DB	218 TQTPSHWSNGHTESIVSENSVIMSVSNRSSSPAGGPGRGLHLGLGGPRD-NSELRH 276
QY	452 ARTEPDSDYRSPHSERY 458
DB	: : : : : : :
DB	277 ARETPDSYRSPHSERH 293
RESULT 6	
Q07112	PRELIMINARY; PRT; 241 AA.
ID	Q07112 PRELIMINARY; PRT; 241 AA.
AC	Q07112;
DT	01-JAN-1998 (TrEMBLrel. 05, Created)
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	Neuregulin-1, glial growth factor 5 isoform precursor (GGFBPP5).
GN	NRG1 OR GGF,
OS	Bos taurus (Bovine).
OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovidae; Bovinae; Bos.
OX	NCBI_TaxID=9913;
RN	[1]
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC	TISSUE-POSTERIOR PITUITARY;
RA	MEDLINE=93205115; PubMed=8096067;
RX	Marchionni M.A., Goodearl A.D.J., Chen M.S., Birmingham-McDonogh O.,
RA	Kirk C., Hendricks M., Danehy F., Misumi D., Sudhalter J.,
RA	Kobayashi K., Wroblewski D., Lynch C., Baldassarre M., Hiles I.,
RA	Davis J.B., Hsuan J.J., rotty N.F., Otsu M., McBurney R.N.,
RA	Waterfield M.D., Stroobant P., Wynne D.;
RT	"glial growth factors are alternatively spliced erbB2 ligands
RT	expressed in the nervous system.";
RL	Nature 362:312-318(1993).
CC	-1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
CC	RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC	RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC	A CTIVATION OF THE ERBB RECEPTORS (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD AND
CC	BRAIN.
CC	-1- ALTERNATIVE PRODUCTS: AT LEAST FIVE DIFFERENT ISOFORMS; GGFBP1,
CC	GGFBP2, GGFBP3, GGFBP4 AND GGFBP5 (SHOWN HERE); ARE PRODUCED
CC	BY ALTERNATIVE SPLICING.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC
 RC TISSUE=POSTERIOR PITUITARY;
 RX MEDLINE=93205115; PubMed=8096067;
 RA Marchionni M.A., Goodearl A.D.J., Chen M.S., Birmingham-McDonogh O.,
 RA Kirk C., Hendricks M., Danehy F., Misumi D., Sudhalter J.,
 RA Kobayashi K., Wroblewski D., Lynch C., Baldassarre M., Hiles I.,
 RA Davis J.B., Hsuan J.J., Totty N.F., Otsu M., McBurney R.N.,
 RA Waterfield M.D., Stroobant P., Gwynne D.;
 RT "glial growth factors are alternatively spliced erbB2 ligands
 RT expressed in the nervous system.";
 RL Nature 362:312-318(1993).
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
 CC ACTIVATION OF THE ERBB RECEPTORS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD AND
 CC BRAIN.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST FIVE DIFFERENT ISOFORMS; GGFBBP1,
 CC GGFBBP2, GGFBBP3, GGFBBP4 AND GGFBBP5 (SHOWN HERE); ARE PRODUCED
 CC BY ALTERNATIVE SPLICING.


```
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00001; EGF_like; 1.
DR SMART: SM00409; Ig; 1.
DR SMART: SM00408; IGC2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
KW Immunoglobulin domain.
FT NON_TER 342 342
SQ SEQUENCE 342 AA; 37836 MW; 8BE36FC836553124 CRC64;

Query Match 26.8%; Score 929; DB 11; Length 342;
Best Local Similarity 70.5%; Pred. No. 3.6e-56;
Matches 203; Conservative 6; Mismatches 48; Indels 31; Gaps 5;

4 PORGRSLSPSRDLFPNP-IRALGPNSPAPRAVVRVRSVSGENSERKEGRKGKGR 62
II II II II II II II II II II II II II II II II II II II II
61 PSGRLLKEDSRKYIFFMEPDANSSGRAPPAPRA-----SPPLETGRNL----- 103

63 GSGKKPESAGSOSPAPLPRLKEMKQESAGSKLVLRCTSSSEYSLRPFKFKNGNELN 122
II II II II II II II II II II II II II II II II II II II II
104 ---KKEYSRLVLCRCALPPLRLKEMKQESAGSKLVLRCTSSSEYSLRPFKFKNGNELN 160
II II II II II II II II II II II II II II II II II II II II
123 RKNKPONIKQKPGKSELINKASLADSGEYMKVSKLGNDSASANTIVNESNIITG 182
II II II II II II II II II II II II II II II II II II II II
161 RKNPENIKQKPGKSELINKASLADSGEYMKVSKLGNDSASANTIVNESNIITG 220
II II II II II II II II II II II II II II II II II II II II
183 MPASTEGAYVSSPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTFVNGGECFM 242
II II II II II II II II II II II II II II II II II II II II
221 MPASTETAYVSSPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTFVNGGECFM 280
II II II II II II II II II II II II II II II II II II II II
243 VKDLSNPSRYLCKCPGFGTCARTENVPV-----KVQNEKAE 281
II II II II II II II II II II II II II II II II II II II II
281 VKDLSNPSRYLCKCPNEFTGDRG-ONYVMASFYMTSRKRKQETKPLE 327
II II II II II II II II II II II II II II II II II II II II

RESULT 9
Q9ESA2 PRELIMINARY; PRT; 323 AA.
AC Q9ESA2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Glial growth factor GGF beta 3 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,
RA Frohner P.W.;
RT "Structural and Functional Diversity of Glial Growth Factor Isoforms
RT Expressed in Regenerating Peripheral Nerve and Associated Neurons.";
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF194996; AAG28450.1; -
DR HSSP: Q12784; 1HRE.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00001; EGF_like; 1.
```

```
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
KW Immunoglobulin domain.
FT NON_TER 1
SQ SEQUENCE 323 AA; 35358 MW; C7DF153A939A80C8 CRC64;

Query Match 26.5%; Score 921; DB 11; Length 323;
Best Local Similarity 65.6%; Pred. No. 1.2e-55;
Matches 198; Conservative 12; Mismatches 38; Indels 54; Gaps 5;

12 PSRDKLPNPPIRALGPNSPAPRAVVRV-----SVSGENSERKEGRK-----GR 56
II II II II II II II II II II II II II II II II II II II II
16 PSTDQ-----PGDPAPYLKVHVQWAVKAGGLKDDSLTDLTGWHPAFPCGR 65
II II II II II II II II II II II II II II II II II II II II
57 GKKEK-----GSGKKPESAGSOSP-----ALPPLKEMK 87
II II II II II II II II II II II II II II II II II II II II
66 LKEDSYIFFMEPDANSSGRAPPAPRASPPPLETGRNLKKEYSVLCKRCALPPLKEMK 125
II II II II II II II II II II II II II II II II II II II II
88 SQESAAGSKLVLRCTSSSEYSLRPFKFKNGNELNKNKPNKIKQKPGKSELINKAS 147
II II II II II II II II II II II II II II II II II II II II
126 SQESAAGSKLVLRCTSSSEYSLRPFKFKNGNELNKNKPNKIKQKPGKSELINKAS 185
II II II II II II II II II II II II II II II II II II II II
148 LADSGEYMKVSKLGNDSASANTIVNESNIITGMPASTEGAYVSSPIRISVSTEGA 207
II II II II II II II II II II II II II II II II II II II II
186 PADSGEYMKVSKLGNDSASANTIVNESNEFITGMPASTETAYVSSPIRISVSTEGA 245
II II II II II II II II II II II II II II II II II II II II
208 NTSSSTSTSTGTSHLVKCAEKEKTFVNGGECFMVKDLSNPSRYLCKCPGFGTCART 267
II II II II II II II II II II II II II II II II II II II II
246 NTSSSTSTSTGTSHLVKCAEKEKTFVNGGECFTVKDLSNPSRYLCKCPNEFTGDR 305
II II II II II II II II II II II II II II II II II II II II
268 NV 269
306 YV 307

RESULT 10
Q96IB3 PRELIMINARY; PRT; 296 AA.
AC Q96IB3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Neuregulin 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC007675; AAH07675.1; -.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
SQ SEQUENCE 296 AA; 31699 MW; 0F5C48C8465B6649 CRC64;

Query Match 7.9%; Score 273; DB 4; Length 296;
Best Local Similarity 42.5%; Pred. No. 4.4e-11;
Matches 62; Conservative 17; Mismatches 39; Indels 28; Gaps 4;

151 SGEYMKVSKLGNDSASANTIVNESNIITGMPASTEG-----AVVSESP----- 197
II II II II II II II II II II II II II II II II II II II II
136 SSEAYTSPVSRQSES-EVQVTVQGDKAVVSPSAAPTKNRFASFSLPSTAPSPSP 194
II II II II II II II II II II II II II II II II II II II II
198 -----IRISVSTEGANTSS-----STSTSTGTSHLVKCAEKEKTFVNGGECFM 243
II II II II II II II II II II II II II II II II II II II II
195 TRNPEVTPKSAIQPOTTETNLTQAPKLSTSTSTGTSHLVKCAEKEKTFVNGGECFM 254
II II II II II II II II II II II II II II II II II II II II
```



```

QY 244 KDLNPSRYLCKCPGFTGARTENV 269
DB 255 KDLNPSRYLCKCPNEFTGDRCONV 280

RESULT 11
Q9ESA8 PRELIMINARY; PRT; 111 AA.
AC Q9ESA8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE SMDF neuregulin beta 2 (Fragment).
GN NRGI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RC TISSUE-AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA Carroll S.L., Anderson K.D., Frohnert P.W.;
RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF194441; AAG28430.1; -.
DR HSSP; Q12784; 1HRE.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 111
FT NON_TER 111
SQ SEQUENCE 111 AA; 12198 MW; CC8BB70584C9F8C CRC64;

Query Match 7.8%; Score 269.5; DB 11; Length 111;
Best Local Similarity 74.6%; Pred. No. 2.2e-11;
Matches 53; Conservative 4; Mismatches 11; Indels 3; Gaps 1;

QY 212 STSTTTGTHLVKCAEKEKFCVNGGECFVKDLNPSRYLCKCPGFTGARTENVPM 271
DB 44 STTSTTSHLIKAEKEKFCVNGGECFVKDLNPSRYLCKCPNEFTGDRCONVY-- 101

QY 272 KVQNOEKAEEL 282
DB 102 -MASFYKAEEL 111

RESULT 12
Q9ESA6 PRELIMINARY; PRT; 256 AA.
AC Q9ESA6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SMDF neuregulin beta 3 (Fragment).
GN NRGI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RC TISSUE-AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA Carroll S.L., Anderson K.D., Frohnert P.W.;
RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AF194443; AAG28432.1; -.
DR HSSP; Q12784; 1HRE.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00001; EGF-like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 256 AA; 27335 MW; C8C08DEA68D33E39 CRC64;

Query Match 7.8%; Score 269; DB 11; Length 256;
Best Local Similarity 41.8%; Pred. No. 6.9e-11;
Matches 61; Conservative 11; Mismatches 46; Indels 28; Gaps 3;

QY 151 SGEYMKYVISKLGNDASANITIVESNIIITGMPA-----STEGAYVSS 194
DB 96 SSEAYTSPVSR-AQSEAGAHVTVQGDHAAVASEPSAVPTRKNRLSAFPFHSAPPFPSP 154

QY 195 ESPRISVSTEGANTSS-----STSTTTGTHLVKCAEKEKTCFVNGGECFV 243
DB 155 ARTPEVTRPKSGTQPTTETNLQAPKLSTSTTGTSHLIKAEKEKTCFVNGGECFV 214

QY 244 KDLNPSRYLCKCPGFTGARTENV 269
DB 215 KDLNPSRYLCKCPNEFTGDRCONV 240

RESULT 13
Q9ESA7 PRELIMINARY; PRT; 136 AA.
AC Q9ESA7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SMDF neuregulin beta 4 (Fragment).
GN NRGI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RC TISSUE-AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA Carroll S.L., Anderson K.D., Frohnert P.W.;
RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF194442; AAG28431.1; -.
DR HSSP; Q12784; 1HRE.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00001; EGF-like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 136
FT NON_TER 136
SQ SEQUENCE 136 AA; 15235 MW; 116CBC91D6E8AEF5 CRC64;

Query Match 7.7%; Score 266; DB 11; Length 136;
Best Local Similarity 69.6%; Pred. No. 4.9e-11;
Matches 55; Conservative 2; Mismatches 12; Indels 10; Gaps 2;

QY 212 STSTTTGTHLVKCAEKEKFCVNGGECFVKDLNPSRYLCKCPGFTGARTENVPM 271
DB 44 STTSTTSHLIKAEKEKFCVNGGECFVKDLNPSRYLCKCPNEFTGDRCONV 102

QY 272 -----KVQNOEKAE 281
DB 103 ASFYMTSRKKQETEKPLE 121

RESULT 14

```

Q8S55		PRELIMINARY;	PRT; 1514 AA.
ID Q8Y55			
AC Q8Y55;			
DT 01-JUN-2002 (TREMBLrel. 21, Created)			
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)			
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)			
DE GH09355P.			
GN G6004.			
OS Drosophila melanogaster (Fruit fly).			
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC Ephydroidea; Drosophilidae; Drosophilla.			
OX NCBI_Taxid=7227;			
RP [1]			
RN SEQUENCE FROM N.A.			
RC STRAIN-BERKELEY:			
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,			
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,			
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,			
RA Miranda A., Mungall C.J., Nunco J., Paclic J., Paragas V., Park S.,			
Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,			
Celniker S.;			
RL Submitted (JAN-2002) to the ENBL/GenBank/DDBJ databases.			
DR ENBL; AY075323; AAL68190.1; -			
SQ SEQUENCE 1514 AA; 157140 MW; 1FTC4B0664105AD5 CRC64;			
Query Match 5.5%; Score 192; DB 5; Length 1514;			
Best Local Similarity 19.3%; Pred. No. 0.00014;			
Matches 135; Conservative 90; Mismatches 298; Indels 176; Gaps 20			
Qy 69 ESAAGOSP-----	:	- - - - -	ALPRLKMKQSQAAGSKLVLCRSTSEYSLSLR 111
Db 272 ESSSSSPVNEPSTGATDSSSTESLPD----	: :	- - - - -	STQESSSSSPVSFFELSTEAT-- 323
Qy 112 FKFKNGNELNNKKPNQIKPKPGSELINKASLADSGEYMCKVKISK-----	: :	- - - - -	LGNDS 166
Db 324 -----	: :	- - - - -	NESSSSSLNSQDSSSSTETSFQTESTDAYDESSTESQDPSTQEISS 376
Qy 167 ASANITIVESNIITGMPASTEGAY-----	: :	- - - - -	VSSSPIRISVTEGANTSSSTSTWT 218
Db 377 STEGPLSTESTAVTDSSSTESSQDSTTQBSSSTEGPLSTESTATENSSSTESSQD 436	: :	- - - - -	
Qy 219 GTSHLVCARKEKTFVNGECFMVKOLSNRYLKCPQGFTGARCTENPMKVQNOEK 278	: :	- - - - -	
Db 437 STT-----	: :	- - - - -	QLSTES-----STEATNESS 463
Qy 279 AEELYQRVLTTICIALVVGMCVVACKTKKQKKLHDLRLQSLRSERNMMNTAN 338	: :	- - - - -	
464 TESSQDST-----	: :	- - - - -	TQESSSTEGPLSTESTATESSTE 500
Qy 339 GPHPNPNNPVNLVNVYKNVISSEHIVERAEAFSTSHYTSTAHHSTVTQTPTP--- 395	: :	- - - - -	
Db 501 SSQ-----	: :	- - - - -	DSTQESSSSSEGPLSTESTATEWNESSSTESSQDSTQEASSSTESPUSLT 554
Qy 396 -----	: :	- - - - -	SHWSNGHTSESILSEHSVIV---MGSVENSR-----HSSP 428
Db 555 EPSTEANESSSTESSQDSTTQEASSSTEDPLSTESTATEWNESSSTESSQDSTTQEASSS 614	: :	- - - - -	
Qy 429 TGGPRGLNCTGGPRECNS-----	: :	- - - - -	FLRHARTPDPSYRDSPHSEYRVVSAMTPARMSP 480
Db 615 TEGLPSTESTEGSNSSSTESSQDSTTQKSSSSTESPLSTEPSTEANESSSTESSQDST 674	: :	- - - - -	
Qy 481 VDFTPSPSKSP-----	: :	- - - - -	PSEMSPPVSMVMSPMAVPFMEERPLLLVTPPLREKKF- 535
Db 675 TQ-ESSSTEGPLSTESTATEWNESSSTESSQDSTTQEASSSSEGPLSTESTATESNESST 733	: :	- - - - -	
Qy 536 ----DHHPQQFSFHHPAHSDNSLPASPLRIVEDEEYETQ-----YEPAQEP 581	: :	- - - - -	
Db 734 ESSQDSTTQEASSSTESPLSTEPSTEANESSSTESSQDSTTQEASSSTEGPLSTEPSTEA 793	: :	- - - - -	
Qy 582 VKKLA-----	: :	- - - - -	NERRAKRTFPNGHANRLEVDNSTSQSSNSESETDERVGEDTPFLGI 635
Db 794 NESSSTESSQDSTTQEASSSSEGPLSTESTATEWNESSSTESSQDSTTQEASSSTEDPLST 853	: :	- - - - -	